5-2016

# Fluorescent Probes to Investigate Homologous Recombination Dynamics

Eric Parker Davenport
*Utah State University*

# FLUORESCENT PROBES TO INVESTIGATE HOMOLOGOUS RECOMBINATION DYNAMICS

by

Eric Parker Davenport

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Biochemistry

Approved:

| | |
|---|---|
| _____<br>Edwin Antony, Ph.D.<br>Major Professor | _____<br>Nicholas Dickenson, Ph.D.<br>Committee Member |
| _____<br>Roger Coulombe, Ph.D.<br>Committee Member | _____<br>Mark McLellan, Ph.D.<br>Vice President for Research and Dean of the School of Graduate Studies |

UTAH STATE UNIVERSITY
Logan, Utah

2016

# ABSTRACT

Fluorescent Probes to Investigate Homologous

Recombination Dynamics

by

Eric Parker Davenport, Master of Science

Utah State University, 2016

Major Professor: Dr. Edwin Antony
Department: Chemistry and Biochemistry

There are multiple mechanisms by which DNA can become damaged. Such damage must be repaired for the cell to avoid ill-health consequences. Homologous recombination (HR) is a means of repairing one specific type of damage, a double-strand break (DSB). This complex pathway includes the Rad51-DNA nucleoprotein filament as its primary machinery. Current methodology for studying HR proteins includes the use of fluorescently labeled DNA to probe for HR dynamics. This technique limits the number of proteins that can be involved in experimentation, and often only works as an end reporter. The work here aims at improving upon standard techniques by creating two fluorescent protein probes. The first probe was developed by directly attaching a fluorophore to *Saccharomyces cerevisiae* Rad51 with the use of click chemistry and the incorporation of unnatural amino acids. This probe could function as a primary reporter on the formation and dissociation of the Rad51-DNA filament in

the presence of pro- and anti- HR mediator proteins. The second probe was created by labeling the exterior cysteine residues of *Plasmodium falciparum* single strand DNA binding protein (SSB) with a fluorophore via maleimide chemistry. This probe acts as a secondary reporter for HR dynamics by signaling for when free single stranded DNA (ssDNA) is available.

(105 pages)

# PUBLIC ABSTRACT

## Fluorescent Probes to Investigate Homologous Recombination Dynamics

Eric Parker Davenport

As the average cell's DNA undergoes roughly 1 million molecular lesions per day, learning about one of the repair mechanisms which fixes damaged DNA is a worthwhile endeavor. Such work may provide insights into how to treat or prevent diseases such as cancers and associated genetic disorders. Homologous recombination (HR) is one pathway which is used by cells to repair a specific type of DNA damage, double-strand breaks. While HR is a complex process involving over 25 different proteins, exactly how HR works, including its regulation, is largely unknown. One protein in particular, Rad51, forms a filament on the damaged DNA and is critical for the repair process.

While current research often utilizes labeled DNA as a means of studying HR proteins, this work involves the development of two alternative means of studying HR. First of all, the Rad51 protein was itself tagged with a fluorophore with the use of a technique developed by Dr. Peter Schultz, which involves altering one of a protein's amino acid residues to a synthetic, unnatural amino acid. The unnatural amino acid incorporated into Rad51 is then labeled with the addition of dye. This results in Rad51 being able to report through fluorescence that it is forming a filament with DNA. A second probe was made by taking single

strand binding protein, which binds to single stranded DNA (ssDNA), and labeling the amino acid cysteine with a fluorophore. This probe works in studying HR by signaling when ssDNA becomes available.

# ACKNOWLEDGMENTS

I would like to thank Dr. Ryan Mehl for providing our lab with the synthetic tRNA and synthetic aminoacyl tRNA synthetase vector. Additionally, I would like to thank Dr. Tom Chang and Jaya Shrestha for assistance in the synthesis of 4-azido-l-phenylalanine. I would like to thank Dr. Edwin Antony for his instruction, mentorship, and support in perusing this work, as well as my committee members Dr. Nicholas Dickenson and Dr. Roger Coulombe. I would like to acknowledge and thank the National Institute of Health for providing the funding, in part, for this research.

Finally, I would like to give a special thank you to the fellow Antony Lab members who worked with me on this project, as well as to my wife and family for their patience and support.

Eric Parker Davenport

CONTENTS

Page

# LIST OF TABLES

# LIST OF FIGURES

# CHAPTER I

# INTRODUCTION

## 1.1 Abstract

*Genetic material is essential for all organisms. Damage to DNA, if not adequately repaired, can have severe and lasting health consequences. One type of damage which DNA may undergo is a double-strand break (DSB). Homologous Recombination (HR) is a pathway for fixing DSB's. To repair the damage, a homologous gene is found and aligned with the DSB. Here the homologous genetic material is used as a template for the addition of the missing nucleotide sequence. Multiple proteins play a role in HR, and notably the Rad51 protein works as the primary machinery in repair. All life utilizes HR to repair damaged DNA. HR is a complex process and is split into the three stages of pre-synapsis, synapsis and post-synapsis. The development of fluorescently labeled protein probes may be useful in the further study of HR and its mediators.*

## 1.2 DNA Damage and HR

Living organisms exist as dictated by a specific genetic code. This genetic code is constructed of deoxyribonucleic acid (DNA). DNA consists of two strands, each made up of many nucleotides. Each nucleotide contains a base (cytosine, thymine, adenine or guanine), attached to a deoxyribose sugar, which is itself attached to a phosphate at the 5′ carbon. Sequential nucleotides are bound together through a phosphodiester linkage, connecting to the 3′ carbon on one sugar and to the 5′ carbon on the next. Due to the complementarity from

hydrogen bonding of the bases cytosine with guanine and thymine with adenine, DNA forms a double helix[1] where each strand forms a complement to the other, one with 3′ to 5′ directionality and the other with a corresponding 5′ to 3′ directionality.

A genetic code provides the information needed to build and maintain an entire organism. In the building of a protein, a sequence of three nucleotides codes for one specific amino acid, with the next three nucleotides coding for the adjacent amino acid in a polypeptide chain. While the instructions for the order of amino acids in proteins are more specifically laid out, the instructions for some structures, such as fatty acids, are put forth by coding for the proteins which synthesize them, such as with fatty acid synthase.

In addition to its structure, the genetic code dictates how a system functions. This can be seen in how DNA codes for signals that travel to other cells, as well as codes for the proteins involved in the sending of damaged or unnecessary proteins to the proteasome to be degraded back into basic building blocks. Also, the genetic code is vital to the survival of an organism's progeny as the next generation's DNA is inherited from the parents. This enhances the importance of DNA's stability, as it must be able to replicate over and over in order to sustain life.

While the stability of DNA remains crucial, it can still be damaged in a variety of ways. Factors such as alkylating agents, mutagens from the environment, electromagnetic radiation, errors in replication, and oxygen radicals

can cause significant damage to nucleotide filaments[2]. Possible resulting damages include single stranded DNA (ssDNA) breaks, base mismatches, insertion, deletions, pyrimidine dimers, bulky adducts, abasic sites, and chemically altered bases[2]. A variety of damages call for a diverse group of mechanisms for DNA repair. These pathways include direct repair, base-excision repair, nucleotide excision repair, and mismatch repair[2]. Fortunately, there is often a code readily available to provide instruction during the repair process. As long as damage occurs to only one strand of the double stranded DNA (dsDNA), the second, undamaged strand can be used as a template for repair.

DNA can undergo an additional type of damage where both strands break apart completely. This is referred to as a double-strand break (DSB). While the two sides of the DSB can be ligated back together, which occurs during nonhomologous end joining (NHEJ), this obvious approach can result in mutations[3] as multiple nucleotides are often lost from either side of the DSB as a consequence of the initial, break-causing event. This can result in a wide range of effects depending on where exactly in the genome the DSB occurs. There is a way, however, to fill in this gap with a sequence of nucleotides that should be harmonious with the good health of the cell. In a diploid cell, every chromosome has a homologous chromosome to pair with it. By aligning the broken chromosome with the still intact one, the gene from the undamaged chromosome can act as a template for the missing nucleotides in a process called homologous

recombination (HR). Other homologous DNA sequences may be found in the genome to be utilized for HR.

All species undergo homologous recombination as a means of DNA repair[4]. HR is a very complex pathway, involving over 30 different genes[5]. The sheer size of a nucleus alone, in relation to the size of just a few missing nucleotides, creates an enormous obstacle that is successfully overcome in the alignment of two homologous sequences.

HR takes advantage to NHEJ as the lost nucleotide sequences at the break are replaced with a nucleotide sequence found in a homologous segment of DNA. Impaired HR function has been linked to diseases, such as breast cancer and ataxia telangiectasia[6]. HR is also seen as an important factor by increasing the amount of genetic diversity during meiosis, creating the necessary linkage between two homologous chromosomes before the initial meiotic division[5]. HR is used to repair single strand gaps that have resulted from DNA damaging agents, the processing of spontaneous damage, or from the collapse of a replication fork, and is additionally responsible for proliferation when telomeres are absent[5].

Homologous Recombination is encompassed by three different stages: pre-synapsis, synapsis and post-synapsis. Figure 1 provides a brief overview of HR. After the occurrence of a DSB, the first event during pre-synapsis is the resection of both ends of the break to form 3′ single stranded overhangs. This is performed by specialized nucleases.

Although homologous recombination is complex and involves over 25 proteins[7], one specific protein appears to be especially central to its execution. A recombination protein, termed Rad51 in yeast, forms a nucleoprotein filament along the 3′ single stranded tails and the adjacent double stranded DNA (dsDNA) (Fig. 2). The 3′ overhangs can be as short as 2 nucleotides in length, and must be complimentary to the homologous DNA[8]. The Rad51-DNA filament drives the search for the homologous gene in the cell, and then works in the alignment of the break with said gene. Rad51 also functions as a catalyst in the strand exchange reaction between homologous sequences of DNA[9].

During the next stage, synapsis, the Rad51-DNA filament performs strand invasion with the homologous gene. This results in D-loop formation (Fig. 1d and Fig. 2c) between the broken and healthy DNA filaments. From here the process shifts to the third, post-synapsis stage, and two distinct pathways are possible, one of which is synthesis- dependent strand annealing (SDSA). Here, following the synthesis of the missing sequence, the strand that initially invades the healthy DNA is displaced and anneals with the 3′ overhang of the other broken strand.[3–5,10]. SDSA is promoted by various antirecombinases disrupting the D-loop[11]. The series of invasion, synthesis, and disengagement can occur repeatedly in the repair of the DSB[4]. SDSA results in no crossover, or the transfer of genetic material, between the two homologous genes. The template gene remains the same as it was prior to the SDSA, and the only change in

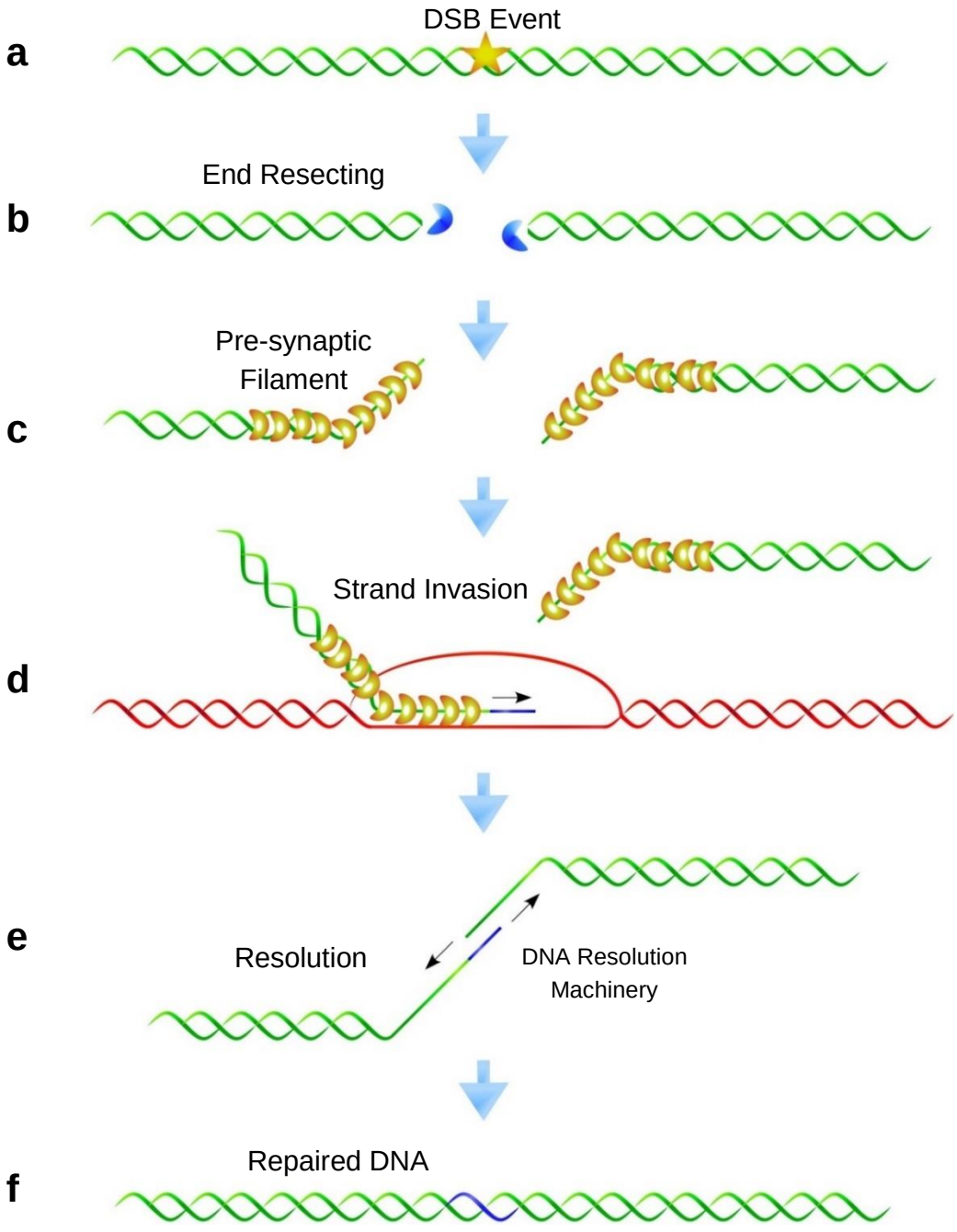



**Figure 1 | A simple model of Homologous Recombination. a,** Some DNA-damaging event occurs, resulting in a double stranded break. **b,** A nuclease resects the two damaged ends, creating 3′ overhangs. **c,** Rad51 filaments form on the resected ends. **d,** After alignment, the homologous gene forms a D-loop and one Rad51-DNA filament performs strand invasion. Synthesis of the missing sequence ensues. **e,** The 3′ overhangs can anneal and complete the necessary synthesis. **f,** The repaired, doubled stranded DNA.

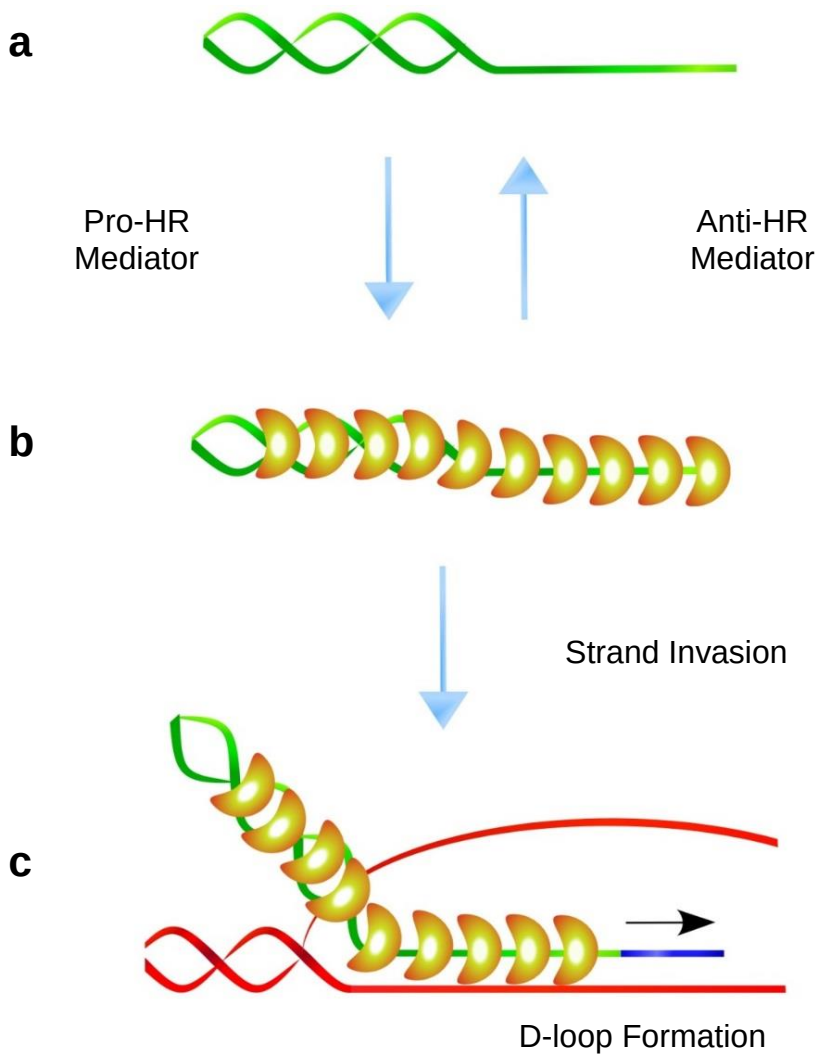



**Figure 2 | The Rad51 filament on a 3′ overhang. a,** Portrayed in green is a resected end of a DSB with a 3′ overhang. **b,** The Rad51 filament (orange) eventually forms on the resected DNA. Pro- and anti- HR mediator proteins will promote either Rad51 filament formation or disassembly. **c,** Rad51 works to align the broken end with the corresponding sequence on the homologous gene (red). Afterwards the missing sequence (blue) is synthesized.

sequence of the repaired strand are the nucleotides synthesized in order to fix the break.

Another option, opposed to the SDSA pathway, is double strand break repair (DSBR). Here the second 3′ overhang also engages with the template DNA, which results in two Holliday junctions. This double junction can be resolved by branch migration, and the dissolution of a hemicatenane by topoisomerase III[4,10]. The process creates non-crossover products and the healthy gene remains unchanged from before to after recombination. In an alternative pathway, still under DSBR, a the double Holliday junction is resolved by an endonuclease[4]. The resulting genes may be either non-crossover or crossover. Crossover products result in a loss of heterozygosity, and appear to occur less commonly in mitotic HR[12]. Resolvases such as RuvABC may limit crossover products through directional cleavage of Holiday junctions.[13] Other, hypothetical pathways beside SDSA and DSBR exist, such as break-induced replication (BIR). Here the amount of nucleotide copied from the healthy chromosome increases as the D-loop is converted to a replication fork[4]. This would result in no crossover between the two chromosomes.

### 1.3 HR Regulation

The proper regulation of HR is critical for cell health. One way to control HR is by regulating Rad51. Since the Rad51-DNA filament is so central to the process of HR, regulating this filament can act as a way to control the level of HR in the cell.

One example of an HR promoter is Rad52. Rad52 binds to Rad51 and is necessary for the nucleation of Rad51 on DNA coated with RPA[14]. Once Rad51

nucleates, it begins to polymerize, displacing RPA. Rad55-Rad57 is a heterodimer and also an HR promoter. This complex is capable of loading Rad51 onto RPA coated DNA, and appears to stabilize and protect the Rad51-DNA filament[15]. This is demonstrated by its formation of a co-filament with Rad51, and resulting inhibition of SRS2 anti-recombinase activity[16].

An additional HR promoter is the Shu complex. This complex consists of the proteins Shu1, Shu2, Csm2, and Psy3. While the role of this complex is not clearly understood[15], recent evidence suggests that the Shu complex works together with Rad52 and Rad55-Rad57 to promote Rad51 filament formation on RPA-coated DNA[17].

The Rad54 protein promotes HR by stabilizing the Rad51 filament and stimulating strand invasion[15], thereby promoting the search for a homologous sequence. Rad54 has also been found to dissociate Rad51 from heteroduplex DNA following strand invasion, allowing for DNA synthesis to commense[18].

Another protein responsible for regulating Rad51 in HR is p53. p53 is a tumor suppressor protein which commonly contains mutations in cancer cells. p53 primarily acts as a transcription factor, regulating the expression of various target genes. In the case of HR, p53 binds the promoter of the Rad51 gene and represses Rad51 expression[19]. This appears to be promoted under conditions with double-strand breaks[19]. p53 also regulates HR by directly binding to Rad51, inhibiting its polymerization[20]. Tumor repressor and pro-HR mediator BRCA2

forms a complex with both p53 and Rad51, and appears to inhibit the transcriptional activity of p53[21].

The HR antagonist SRS2 exhibits both anti-recombinase and helicase activity. SRS2 works to disassemble the Rad51-DNA filament[22]. SRS2, as well as helicases Mph1 and Sgs1, appear to favor SDSA over DSBR, thus limiting crossover products[23,24]. Deficiency in the HR promoter BRCA1 has been associated with breast and ovarian cancers[25]. Further research has linked defective HR with methylation at the BRCA1 promoter [26]. Recent evidence also indicates that serine/arginine-rich splicing factor 3 (SRSF3) regulates the expression of BRCA1 and RAD51 genes [27]. Further proteins not detailed here also function in mediating homologous recombination.

If an inappropriate recombination event occurs, the genomic stability becomes threatened. When DSB's are not correctly or fully repaired, carcinogenesis, chromosome loss, chromosomal rearrangements, and apoptosis may occur[6]. A current area of interest in HR includes, in addition to the formation and function of the Rad51 filament, how HR regulators effect the different stages of DSB repair[10,28,29].

### 1.4 Rationale for Thesis Research

While the enzymes involved in HR have been identified, the precise mechanism of how they function to coordinate HR is poorly understood. For example, how the cell decides to commit to HR or abort unnecessary HR events

is not known. To understand how Rad51 filaments are formed and disassembled during HR, a real-time tool to quantitate Rad51 dynamics on DNA is required.

Current experimental tools to observe Rad51 dynamics rely on fluorescently labeled DNA. In these experiments the fluorescence of DNA is monitored and when Rad51 binds to the DNA, an increase in fluorescence intensity is observed. Therefore, binding and dissociation of the Rad51-DNA filament can be measured. Such fluorophores are often placed at the end of DNA (and on short DNA oligonucleotides). Hence the assay reports HR events happening only at the ends. The assay is also limited to observing only one protein at a time; hence multi-protein HR reactions cannot be reconstituted. To overcome such limitations our research focuses on developing fluorescent HR proteins which, upon binding to DNA, will display a change in fluorescence. This approach will overcome the end-effect and DNA length limitations associated with current techniques.

**Hypothesis:** Rad51 binds to ssDNA during HR. We hypothesize that regions L1 and L2 of Rad51 will undergo conformational changes upon binding to DNA. By placing fluorophores at these regions, we hypothesize that changes in fluorescence upon ssDNA binding can be visualized. To help test these hypotheses, the following specific aims will be addressed.

**Aim 1**: Create a probe for the Rad51-DNA filament by incorporating the unnatural amino acid 4-amino-l-phenylalanine in place of specific amino acid residues in *Saccharomyces cerevisiae* Rad51.

**Aim 1a**: Specifically investigate positions on loops L1 and L2 of *Sc*Rad51 which, following labeling with an alkyne fluorophore, will result in fluorescently labeled protein which displays a change in fluorescence emission intensity upon filament formation.

**Aim 1b**: Develop a UAA probe with a fluorescent label near the N-terminus of *Sc*Rad51, to be used for future FRET and single molecule designs.

**Aim 2**: With *Plasmodium falciparum* single strand DNA binding protein (SSB), develop a probe which, when labeled with a maleimide fluorophore, will act as a signal for unbound, single stranded DNA (ssDNA) by undergoing a change in fluorescence emission intensity upon binding.

The ability to visualize Rad51 filament formation on, and dissociation from, DNA in real time would be very useful in the study of homologous recombination. This could act as an indicator of what regulator proteins are pro- or anti- Rad51-DNA filament formation, and thus which proteins are pro- or anti- recombinases. While classifying a protein or a system, it is useful to get real-time data, rather

than waiting for reactions to reach equilibrium prior to data being taken. Quick, real-time analysis of molecular interactions not only provides kinetic detail, but can give insights into the reaction mechanism. Such real-time experiments could also function to determine the rate of Rad51 filament formation under differing conditions. Additionally, single molecule studies could be pursued with this technique with the visualization of an individual, labeled Rad51 molecule.

This work focuses on creating probes that allow for the visualization of Rad51 filament dynamics, and focuses specifically on fluorescently labeled proteins. A fluorescent label may increase or decrease in signal intensity depending on the conformation and environment of the protein it's attached to. By focusing on proteins that undergo both conformational and environmental changes, depending on whether a Rad51 nucleoprotein filament forms or dissociates, we can develop probes that have the potential to reveal new insights into the inner workings of homologous recombination and its regulators.

New therapeutic approaches might investigate HR proteins as targets for treatments of genetic disorders and diseases. As Rad51 expression is increased in various cancers, one direction of research for potential therapies has been to deliver a gene containing the Rad51 promotor and an open reading frame containing a gene detrimental to cell health, such as the diphtheria toxin A (DTA) gene, to the patient's cells[30,31]. The probes developed here might help further elucidate the function of Rad51 filament mediator proteins, indicating which proteins/genes could be targeted for therapeutic purposes.

# CHAPTER II

# DEVELOPING A RAD51 PROBE USING UNNATURAL AMINO ACIDS

## 2.1 Abstract

*Formation of the Rad51 filament on single stranded DNA (ssDNA) is one of the key events in the homologous recombination (HR) pathway and is under strict regulation by both pro- and anti- HR mediator proteins. The formation of the Rad51 filament and its disassembly dictate the fate of HR and this study has focused on developing a real-time assay to quantitate Rad51 dynamics on ssDNA. Utilizing unnatural amino acids as an attachment point for fluorophores, fluorescent Rad51 probes have been developed which, when bound to ssDNA, undergo a change in fluorescence. These studies further the understanding of how Rad51 nucleoprotein filaments are formed, and will be useful for future studies aimed at understanding how mediator proteins alter its dynamics. This information is critical in understanding the role of mediator proteins in HR associated defects, and to uncover why mutations in these proteins lead to genomic instability, cancers and associated genetic disorders.*

## 2.2 Introduction

As noted in the previous chapter, Rad51 is the primary machinery behind homologous recombination. Rad51 exists in both the nucleus and the cytoplasm of the cell [32,33]. It plays a role in both finding a homologous sequence and aligning the double-stranded break for repair. Targeting a probe that allows for

the visualization of the nucleoprotein filament formation would be beneficial, as it could help understand how pro- and anti- HR mediators function, as well as provide insights into how homologous chromosome search, alignment, and strand invasion are performed. Taking a closer look at the process of homologous recombination further reveals how a Rad51 probe could be useful in studying the HR pathway.

Following the DSB event, and after the two broken ends are resected, the single-stranded overhangs are bound by Replication protein A (RPA). The RPA protein is a heterotrimer, formed of three subunits: RPA1, RPA2, and RPA3. This protein has multiple roles, functioning in replication, recombination and repair for eukaryotes[34]. As ssDNA is susceptible to damage in the cell, RPA works to protect the nucleotide strand from degradation. RPA also functions as a sensor to signal for the DNA damage response[35]. Additionally, RPA marks the DNA as being single-stranded. As the bound RPA is displaced from ssDNA, the Rad51 filament forms simultaneously[36]. Proteins can be pro-recombinases by promoting the dissociation of RPA or promoting the association of Rad51. Anti-recombinases work in a differing manner, promoting the dissociation of Rad51 from DNA (Fig. 3).

### *2.2.1 Finding the Appropriate Probe Technique*

In the pursuit of studying the formation and dissociation of the Rad51-DNA filament, one utilized probe has been DNA labeled with a 3′ or 5′ fluorophore[37,38]. By attaching a 3′ or 5′ fluorophore to DNA, the formation of a nucleoprotein

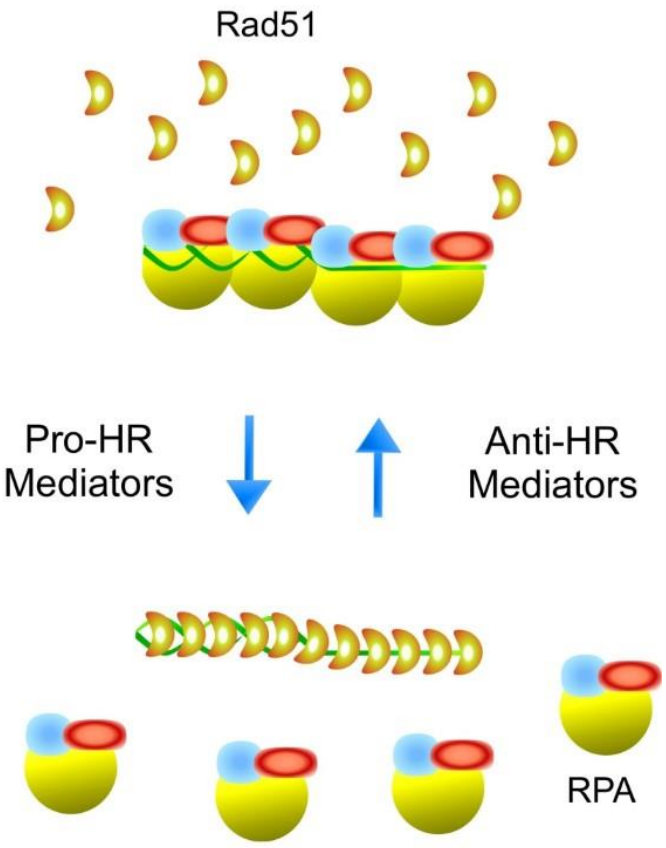



**Figure 3 | RPA and Rad51.** Once the resecting at the ends of a double-strand break has taken place, RPA binds the resulting single-stranded overhangs. Pro-HR mediator proteins can then work to promote the removal of RPA and the formation of a Rad51-DNA filament. This is conducive to HR taking place in the cell. An anti-HR mediator can work antagonistically to this process, assisting in the disassociation of Rad51 and the binding of RPA at the 3′ overhangs.

filament can be visualized by an increase or decrease in fluorescent intensity. This technique, while it has been used for significant contributions, is limited in that the probe is just an end reporter. Additionally, when multiple HR proteins are present in an experiment, questions can arise as to which proteins are causing the observed change in fluorescence signal.

To improve upon this technique, Rad51 becomes a candidate for conversion into a probe for the study of HR dynamics. The creation of a Rad51 probe could provide a direct signal for both the formation and dissociation of the Rad51-DNA filament. This would provide insights into which proteins are promoting or antagonizing HR. In order for Rad51 to have sufficient viability as a probe, some quality of the protein must be quantifiable and measured in a laboratory setting. Additionally, the signal must change to some degree based on environmental, conformational, or spacial changes undergone by the probe.

Conversely, RPA could also be targeted as a probe. However, as Rad51 is the central machinery to the HR process, it has been chosen for this study. Future work in the Dr. Edwin Antony Lab will investigate the utilization of RPA as a probe to study HR.

Quantitating fluorescence of the amino acid tryptophan is a technique available to measure changes in a protein's environment or confirmation. Tryptophan has an aromatic side chain which undergoes fluorescence at an excitation of 280nm and an emission of 348nm. However, quantitating tryptophan fluorescence for yeast Rad51 is not an option as the protein has no tryptophan residues. This technique would also be limited as a current DNA-based assay, as multiple proteins cannot be used at once. When measuring tryptophan fluorescence does not work as an adequate method for obtaining insights into protein dynamics, the addition of a nonnative fluorophore directly to the protein may prove beneficial.

A fluorophore functions by absorbing a photon at a given wavelength, causing an electron to be excited to a higher energy state. While at this higher energy state, the electron may lose some energy by vibrational relaxation. Finally, as a photon is emitted from the fluorophore, the electron relaxes to its ground state. If any energy was dissipated during the exited state, the emitted photon will have lower energy, and a corresponding longer wavelength, than the exciting photon. While labeling a protein with a fluorophore can cause the protein to fluoresce regardless of its activity level, an added benefit of labeling with a fluorophore comes when a change in fluorescence emission intensity is seen as activity occurs. This can come about from a change in the microenvironment of the fluorophore, or from a conformational change in the protein itself.

Multiple concerns can arise when labeling a protein with a fluorophore. While the fluorophore might be placed where some change in fluorescence is observed as the protein undergoes activity, the ideal placement of the fluorophore should not interfere with protein activity or conformation. This can arise as an issue especially when the fluorophore, and whatever is attaching it to the protein, add large and bulky components to the protein's surroundings[39]. If the label is too intrusive, then the fluorophore will not be a reliable reporter for what actually occurs when the fluorophore is absent.

A common technique for labeling proteins with a fluorophore is through maleimide chemistry. The carbon-carbon double bond of a maleimide group can react with a sulfhydryl group, resulting in a stable, nonreversible thioester

linkage. A sulfhydryl group is found in the side chain of the amino acid cysteine. A protein with exterior cysteine residues can hence have sulfhydryl groups available for this reaction. Therefore, if a fluorescent molecule with a maleimide group is mixed with this protein, a reaction can occur linking the protein and fluorophores. This technique is simple, and does not include bulky molecules between the protein and fluorophore to link them together.

Creating a fluorescent Rad51 with maleimide chemistry, however, becomes problematic as the Rad51 protein in *S. cerevisiae* contains 8 cysteine residues. Several of these residues are important for Rad51 function[40,41]. As cysteine-labeling is not a viable option with Rad51, an alternative method must be turned to, although such an approach could be used for other proteins.

### *2.2.2 Unnatural Amino Acid Incorporation Technique*

One recently developed technique for labeling proteins has been to incorporate an unnatural amino acid (UAA) into the polypeptide chain itself[42]. UAA's are amino acid derivatives that are not commonly found in nature. A UAA may contain a reactive group on its side chain. In order to work as a probe, the placement of a UAA must not interfere with the folding, confirmation, or activity of the protein of interest.

The UAA of choice here, 4-azido-L-phenylalanine (4AZP), is a phenylalanine derivative and contains an azide group on the 3′ carbon of the side-chain benzene ring. A method has been developed by the Shultz group for

incorporating UAA's into proteins[43]. Furthermore, there has been much work on expanding the methods and use involving UAA incorporation in proteins[44–47].

The first step in incorporating an unnatural amino acid is having a synthetic tRNA which recognizes a specific codon sequence in the mRNA as a point to insert the UAA. Likewise, in order to charge the synthetic tRNA with a UAA, a synthetic aminoacyl tRNA synthetase must be developed. Dr. Ryan Mehl's Lab at Oregon State University has developed a synthetic tRNA, as well as a synthetic aminoacyl tRNA synthetase, for 4AZP. A pDULE vector coding for both the tRNA and the synthetase have been provided by the Mehl group. The synthetic tRNA synthetase was designed for the specific purpose of charging the synthetic tRNA with 4AZP. The charged tRNA is then capable of reading an amber stop codon in the mRNA, not as a site to terminate translation, but rather as a specific position to insert 4AZP, allowing for the continuation of translation. See Figure 4. It is important to note, however, that the cell still has its natural environment where the stop codon can be recognized as a stop codon. Here the amber stop codon is recognized at the A site of the ribosome by release factor 1 (RF1). RF1 then promotes the release of the polypeptide, and translation is terminated. This results in a truncated protein product, which does not carry the 4AZP.

The next step in the Rad51-UAA synthesis is to decide which codon sequence should be altered to an amber stop codon, causing the translated product to have an UAA instead of the regular amino acid. Selecting the proper

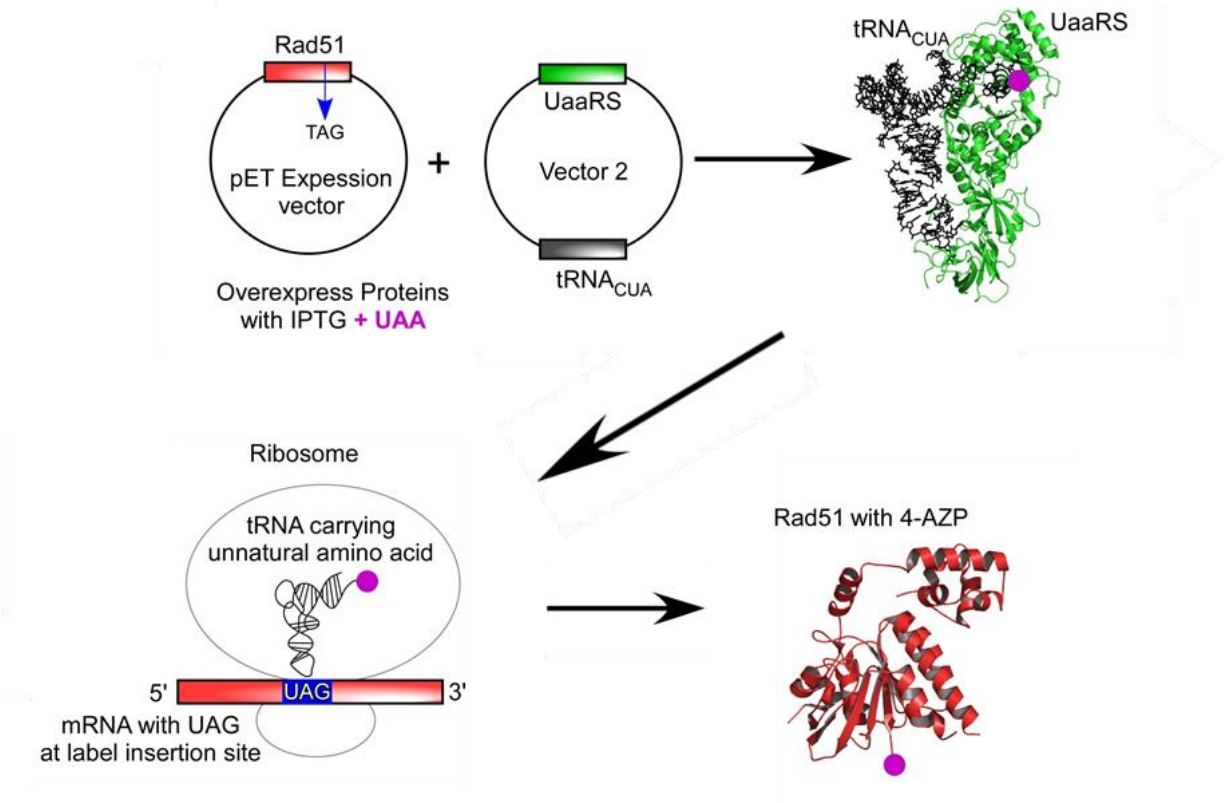



**Figure 4 | Unnatural amino acid incorporation.** In an expression vector coding for the protein, an amber stop codon (TAG) substitution is performed at the residue of choice for 4AZP placement. This plasmid is then co-transformed into protein overexpression cells with a second vector coding for both the synthetic tRNA and unnatural aminoacyl tRNA synthetase (PDB, 4JXZ). At the ribosome, the UAG stop codon is read by the synthetic tRNA, charged with 4AZP (purple), as a point for 4AZP insertion. Upon successful incorporation, translation continues and the resulting Rad51 protein (PDB, 1SZP) contains 4AZP at the desired residue.

site to change is essential in creating a functional probe. Selecting a residue which has a side chain open to the surface allows for the residue to be available for the chemical addition of a fluorophore through click chemistry. Additionally, the selected residue for UAA incorporation needs to be at a position where the attached fluorophore will undergo a change in fluorescence emission intensity

upon the formation or dissociation of a Rad51-DNA filament. If no change in fluorescence is observed when a nucleoprotein filament forms and dissociates, then the probe will not work as a signal for what is happening to the Rad51 filament.

*2.2.3 Selecting Residue Sites for Investigation*

The crystal structure of Rad51 was used to select an appropriate position for UAA incorporation. In addition, a high-resolution structure of the homologous RecA protein, bound to DNA, was used as a guide for design. By comparing crystal structures of RecA[48] and Rad51[49], homology between the proteins is observed (Fig. 5a).

While structures for Rad51 and RecA are both available, only RecA has a solved structure in which it is bound to DNA[48]. When comparing this structure to the crystal structure where RecA is not bound DNA, a conformational change is seen in two loops, deemed loops L1 and L2. Rad51, having high homology to RecA, has corresponding loops which might also undergo a conformational change upon binding to DNA (Fig. 5b). Loops L1 and L2 make good sites for investigation of UAA incorporation as they are likely near the space where Rad51 associates with DNA in the nucleoprotein filament, and could possibly undergo a change in confirmation upon DNA binding, as occurs in the RecA-DNA filament formation.

Correspondingly, two sites where investigated on loop L1 (F290 and S291), along with one site on loop L2 (F337). A phenylalanine residue was

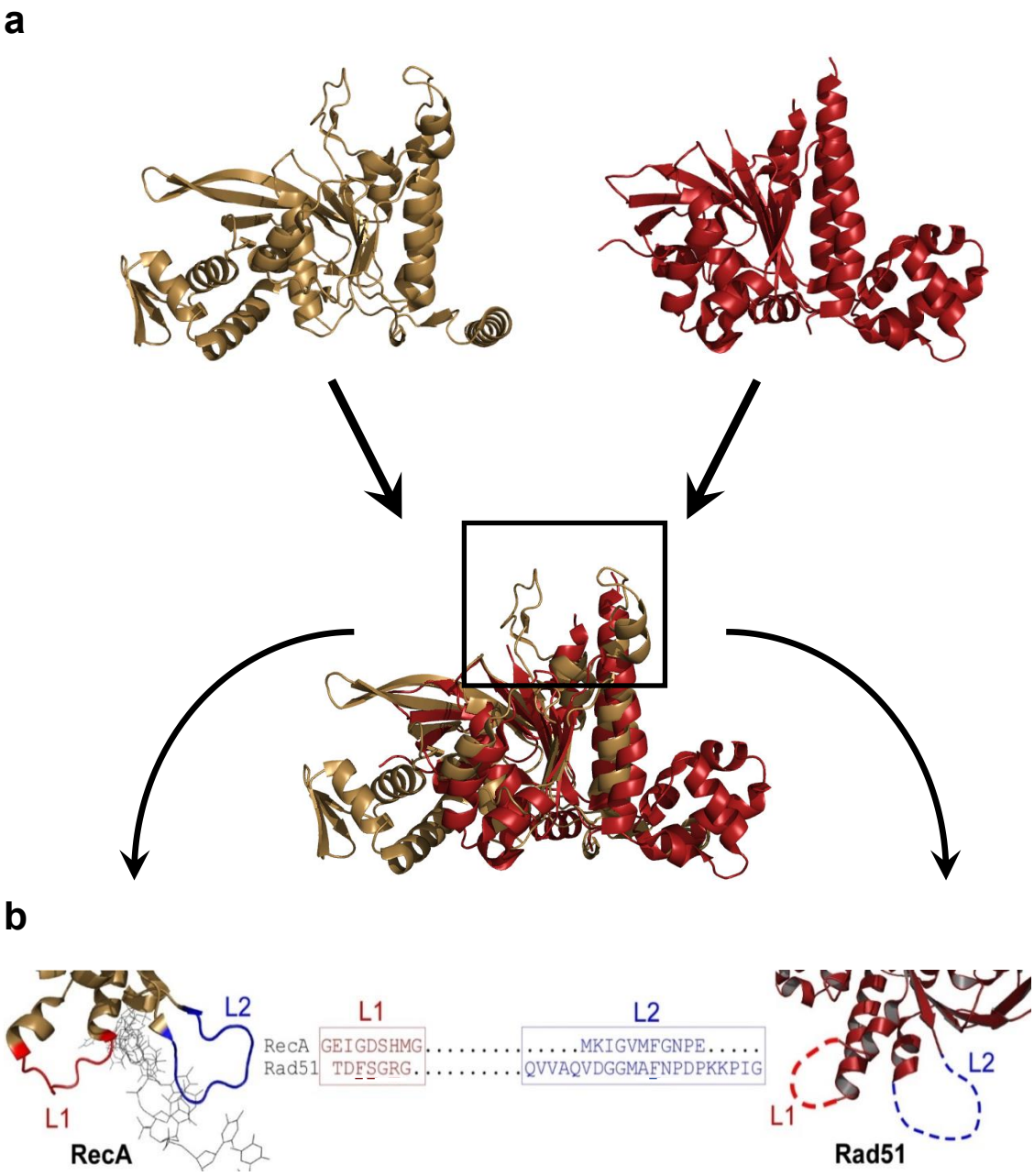



**Figure 5 | Selecting DNA-binding residues. a,** The crystal structures of RecA (brown) and Rad51 (red) are superimposed, highlighting homology. **b,** On the left is the crystal structure for RecA, bound to ssDNA. Loops L1 and L2 undergo a conformational change upon binding. Rad51, on the right, has similar loops L1 and L2, which are disordered when not bound to DNA, but may also undergo a change in confirmation upon binding to DNA. Crystal structure images were generated from PDB 3CMU, 3LDA, and 1SZP.

selected in each loop, as 4-azido-L-phenylalanine resembles phenylalanine. This is favorable, as a smaller change in structure should cause less disruption in protein confirmation and function.

*2.2.4 An Alternative Label for Filament Dynamics*

An additional site near the N terminus of Rad51 was investigated for UAA incorporation (Y16). The N terminal region of *Sc*Rad51 is not seen in human and bacterial homologs (Fig. 6). *Sc*Rad51 has shown the same strand exchange activity with and without the N terminus[50]. Furthermore, the crystal structure of Rad51 has only been solved without the N terminal domain included[49,50]. It may be that the N terminal domain is too floppy to result in crystallization. By selecting a site that is not associated with protein function, the incorporation of the fluorophore may have minimal change on the activity.

This probe, as it is not near the hypothesized DNA binding site, is not as likely to undergo a large change in fluorescence upon filament formation. However, if two appropriate fluorophores are used, fluorescence resonance energy transfer (FRET) may be utilized. Under FRET, after a fluorophore is excited by an incoming photon, it non-radiatively transfers energy to a nearby fluorophore instead of emitting a photon of its own. The second fluorophore is then free to emit a photon upon relaxation. In order for FRET to occur, the emission wavelength for the first fluorophore must be within the spectrum of the excitation wavelength of the second fluorophore. This allows for the proper amount of energy to be transferred between the two molecules. Such is the case

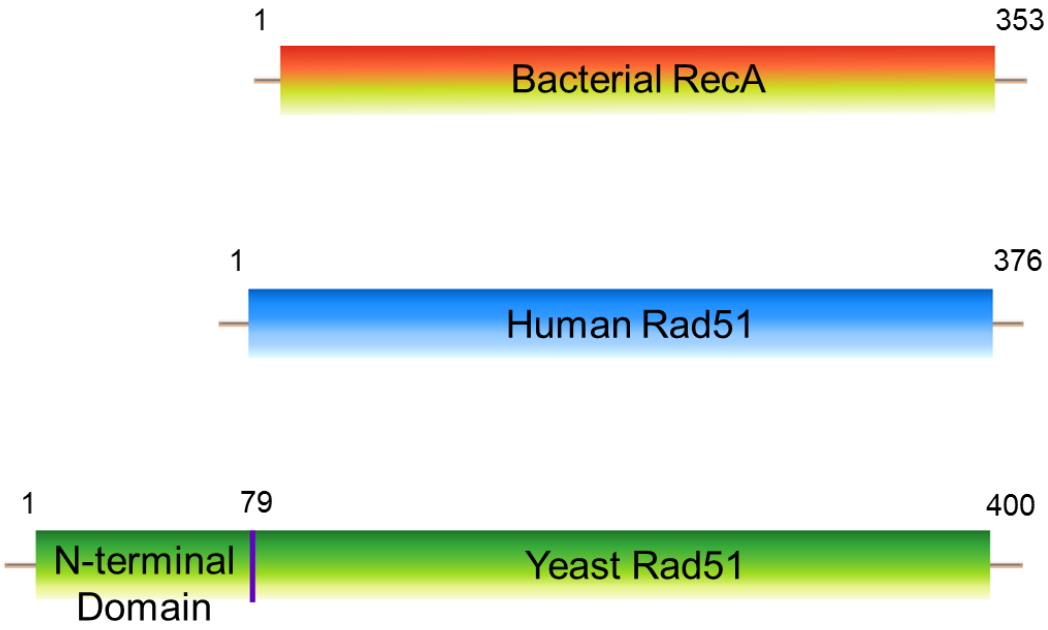



**Figure 6 | Yeast Rad51 has an extra N-terminal domain.** Yeast Rad51 contains an extra N-terminal domain which is not found in both bacterial RecA and Human Rad51.

with the Cy3 and Cy5 fluorophores, and they have been used in previous FRET experimentation[51–53].

The distance between two fluorophores in FRET experiments is driven by their Foster radius, the distance at which energy transfer efficiency is 50%. Previous studies have utilized the Cy3-Cy5 FRET pair with Foster radius values of 54-61 Angstroms[53]. While the crystal structure for *Sc*Rad51 does not include the N terminus[50], the distance between two N termini can be estimated from the solved bacterial RecA filament structure[48], which is 49.9 angstroms (Fig. 7). This distance, being near the Foster radius calculated in previous studies, indicates

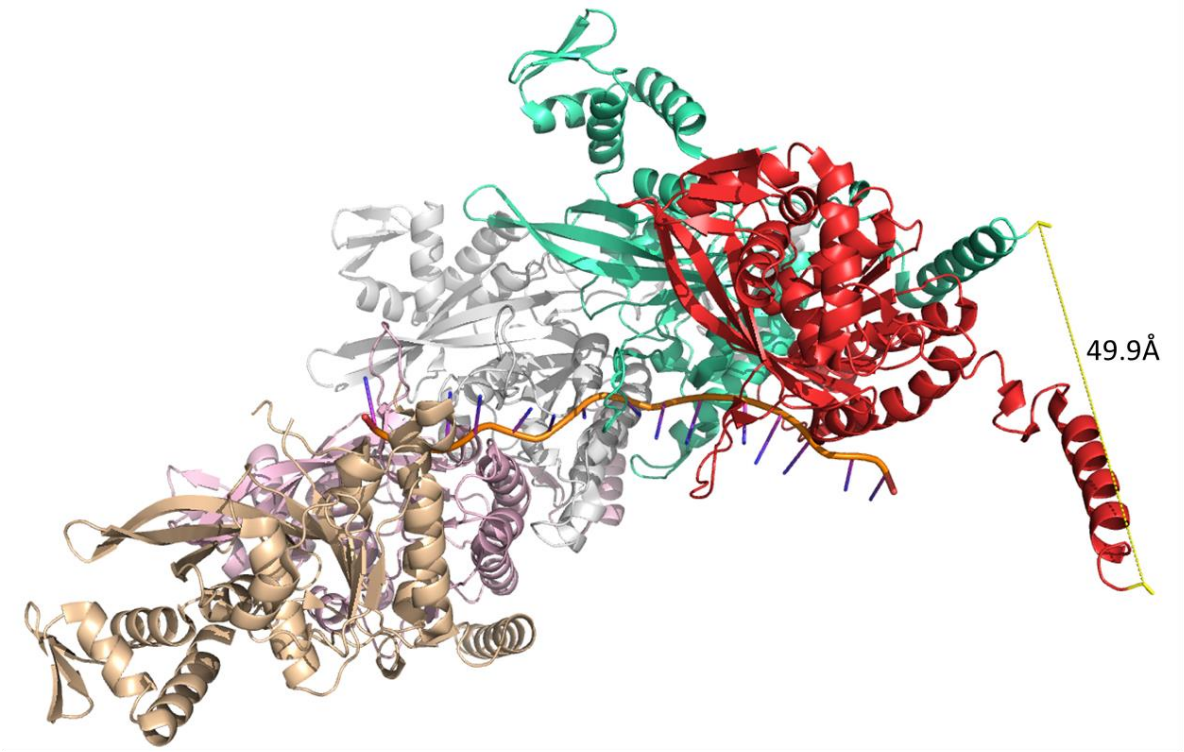



**Figure 7 | The RecA-ssDNA Filament.** The RecA filament (PDB, 3CMU) with monomers in various colors formed on ssDNA (orange). The distance between two adjacent N termini is measured to be 49.9 angstroms, ideal for FRET experimentation.

that FRET experimentation may be a worthwhile endeavor with the 16-UAA probe, as it is positioned near the N terminus of a homologous protein

## 2.3 Materials and Methods

### *2.3.1 Mutations*

The *Sc*Rad51 DNA sequence was incorporated into a pTXB1 vector, using the Nde1 to Nhe1 sites. An EcoR1 site, separating the gene of interest from the chitin binding domain (CBD), was removed and replaced with a single histidine residue, to act as a linker between Rad51 and the CBD. As pTXB1

includes ampicillin resistance, this antibiotic was included in the utilized media when cells containing the plasmid were proliferated.

To alter the genetic coding for the specific amino acids desired to be changed to sites for UAA incorporation, site-directed mutagenesis was performed. For example, the sequence for phenylalanine 337 (F337) was mutated from TTT to TAG (amber stop codon) in this manner. Amber stop codons were substituted into the gene with the use of a Q5® Site-Directed Mutagenesis Kit (New England Biolabs). Primers were developed with the associated NEBBaseChanger™ software (New England Biolabs). All UAA clones contained only one TAG mutation as confirmed by sequencing.

The synthetic aminoacyl tRNA synthetase and the synthetic tRNA were both contained in a pDULE vector. The prepared vector was provided to our lab by the Dr. Ryan Mehl group at Oregon State University.

### *2.3.2 Rad51 Protein Expression*

Rad51 and synthetic tRNA vectors were co-transformed into BL21-AI One Shot® Chemically Competent *E. coli* cells (ThermoFisher Scientific). One colony was selected for a 30 mL starter growth, along with 100 µM ampicillin and 50 µM spectinomycin.

The expression media was specifically prepared for this strategy, and followed a protocol outlined in *Preparation of site-specifically labeled fluorinated proteins for $^{19}$F-NMR structural characterization*[54]. Components of solutions can

be found in Tables 1-4. The expression media was combined as outlined in Table 1.

For a 50x M salt solution, the chemicals listed in Table 2 were combined in 500 mL of MQ $H_2O$. Table 3 includes the components for the 25x Amino Acid Mix, which includes every amino acid save tyrosine and cysteine. Tyrosine has a low solubility (2.5mM) at neutral pH, while cysteines readily convert to cystines in media[55]. Cysteine thiol groups may also generate the reactive oxygen species hydrogen peroxide[56]. Fortunately, *E. coli* is capable of synthesizing both of these amino acids[57]. 5g of each included amino acid was added to the mix, and the solution was filled to 1 L with MQ $H_2O$ for a final concentration of 400 µg/mL. The components of the Trace Metal Solution are listed in Table 4.

The components found in Table 1 were combined with 408 mL of autoclaved dd$H_2O$ in a 2 L Erlenmeyer flask. A second flask, with the same solution, was prepared as a negative 4AZP control. 2.5 mL of overnight seeding culture was added to each flask, along with 500 µL of both 100 mg/mL ampicillin and 50 mg/mL spectinomycin. Each flask was shaken at 250 RPM and 37ºC, as optical density at 600 nm ($OD_{600}$) was monitored with a spectrophotometer up to an OD of 2.0.

To induce protein expression, a final concentration of 400 µM IPTG was added to each flask for lac operon induction of the pTXB1-Rad51 and pDULE-tRNA vectors. Furthermore, 167 µM L-arabinose was added at induction, as the gene encoding T7 RNA polymerase in the BL21-AI *E. coli* cells is controlled by

**Table 1 | UAA expression media**

| Component | Sterilization | Amount |
|---|---|---|
| 25x Amino Acid Mix | | 20 mL |
| 5% Aspartate, pH 7.5 | Autoclave | 25 mL |
| 40% (wt/vol) Glucose | Autoclave | 600 µL |
| 10 % (vol/vol) Glycerol | Autoclave | 25 mL |
| 4 mg/mL Leucine pH 7.5 | Autoclave | 10 mL |
| 1 M $MgSO_4$ | Autoclave | 1 mL |
| 50x M Salts | Autoclave | 10 mL |
| Trace Metal Solution | | 100 µL |
| Water | Autoclave | Fill to 500 mL |

**Table 2 | Salts for UAA expression.**

| Component | 50X Concentration | Amount |
|---|---|---|
| $NH_4Cl$ | 2.5 M | 66.9 g |
| $KH_2PO_4$ | 1.25 M | 85.1 g |
| $Na_2HPO_4$ | 1.25 M | 85.1 g |
| $Na_2SO_4$ | 0.25 M | 17.8 g |
| Water | | Fill to 500 mL |

an *ara*BAD operon. 103 mg of 4-azido-phenylalanine was dissolved in 125 µL of 5 M NaOH, and then brought to a final volume of 4 mL with MQ $H_2O$. All 4 mL of the solution were added to the 4AZP positive control flask. The inclusion of 0.048% (wt/vol) glucose in the expression media helps inhibit leaky plasmid expression. *E. coli* cells will exhibit catabolite repression by consuming available glucose prior to the consumption of lactose, xylose or arabinose[58,59].This occurs as the transcriptional activation of genes involved in catabolizing these lower-priority sugars is prevented[60]. Limiting the consumption of arabinose prior to induction will inhibit the expression of T7 RNA polymerase, and thus repress the expression of the pTXB1-Rad51 and pDULE-tRNA vectors.

**Table 3 | 25X amino acid stock**

| Component | Amount |
|---|---|
| Alanine | 5 g |
| Asparagine | 5 g |
| Aspartic Acid | 5 g |
| Arginine HCl | 5 g |
| Glutamic acid sodium salt | 5 g |
| Glutamine | 5 g |
| Glycine | 5 g |
| Histidine HCl $H_2O$ | 5 g |
| Isoleucine | 5 g |
| Leucine | 5 g |
| Lysine HCl | 5 g |
| Methionine | 5 g |
| Phenylalanine | 5 g |
| Proline | 5 g |
| Serine | 5 g |
| Threonine | 5 g |
| Tryptophan | 5 g |
| Valine | 5 g |
| Water | Fill to 1 L |

The addition of 4-azido-L-phenylalanine allowed for the synthetic tRNA to become charged and to recognize the amber stop codon as a site to add its loaded UAA. When no 4AZP was introduced, the amber stop codon was solely read as a point to terminate translation.

After inducing for three hours at 37ºC and 250 RPM, cell growths were pelleted at 10,000 RPM for 30 minutes. Supernatant was discarded, and cell pellets were stored at -80ºC.

**Table 4 | Trace metal stock**

| Component | Amount for 30 mL Stock | Amount from Stock for 5,000X solution | Final Media Concentration |
|---|---|---|---|
| $CaCl_2$ - $2H_2O$ | 8.82 g | 500 ml | 4 µM |
| $MnCl_2$ - $4H_2O$ | 5.93 g | 500 ml | 2 µM |
| $ZnSO_4$ - $7H_2O$ | 8.62 g | 500 ml | 2 µM |
| $CoCl_2$ - $6H_2O$ | 1.32 g | 500 ml | 400 nM |
| $CuCl_2$ | 807 mg | 500 ml | 400 nM |
| $NiCl_2$ | 777 mg | 500 ml | 400 nM |
| $Na_2MoO_4$ - $2H_2O$ | 1.45 g | 500 ml | 400 nM |
| $Na_2SeO_3$ | 1.03 g | 500 ml | 400 nM |
| $H_3BO_3$ | 371 mg | 500 ml | 400 nM |
| $FeCl_3$ | 486 mg | 25 ml | 10 µM |
| Water | Fill to 30 mL | Fill to 50 mL | |

*2.3.3 Rad51 Protein Purification*

The resuspension of the cell pellet was performed with 50 mL of Resuspension Buffer [0.1 M Tris-HCl, pH 8.0, 5 mM EDTA, pH 8.0, 500 mM NaCl, 1 M urea, 5 mM 2-mercaptoethanol, 10% wt./vol. sucrose, 10% wt./vol. glycerol, and 1 mL Protease Inhibitor Cocktail (2 mM AEBSF, 0.3 µM Aprotinin, 116 µM Bestatin, 14 µM E-64, 1 µM Leupeptin and 1 mM EDTA)] (Sigma-Aldrich). This resuspension was transferred to a beaker, and after the addition of lysozyme (40mg/100mL), the solution was allowed to stir in a cold room for 30

minutes. Afterwards, 6 mL of 100 mM sodium deoxycholate, dissolved in 0.125 M NaOH, was added to the lysate. Following an additional 20 minutes on the stir plate, sonication was performed to ensure complete cell lysis. Sonication was carried out at 50% duty cycle for 1 minute on and 1 minute off. This was repeated, and the resulting lysate was spun down at 17,000 RPM for 60 minutes.

As the soluble protein remained in the supernatant, the pelleted debris were discarded. Next, an ammonium sulfate pellet precipitation was performed. Increasing the ammonium sulfate concentration, and thus the concentration of ions, to high levels will decrease the solubility of proteins[61]. Proteins which aggregate more easily salt-out at lower concentrations of ammonium sulfate than proteins with higher solubility[62].

For this step, the supernatant was transferred to a beaker and placed on a stir plate at 4ºC. 0.24g/mL ammonium sulfate was added to the supernatant slowly over a period of 15 minutes while stirring, forcing the Rad51-CBD fusion protein to precipitate out of the solution (Fig. 8). After an additional 20 minutes, the solution was spun down at 10,000 RPM for 30 minutes.

The ammonium sulfate pellet was resuspended in 25 mL of AS Resuspension Buffer (20 mM Tris-HCl, pH 8, 0.5 M NaCl, 0.1 mM EDTA, 10% glycerol, and 0.1% TritonX100). 10 mL of Chitin Bead Slurry (New England Biolabs) was prepared with two MQ $H_2O$ washes and three AS Resuspension Buffer washes. The 10 mL of Chitin Bead Slurry was then combined with the 25

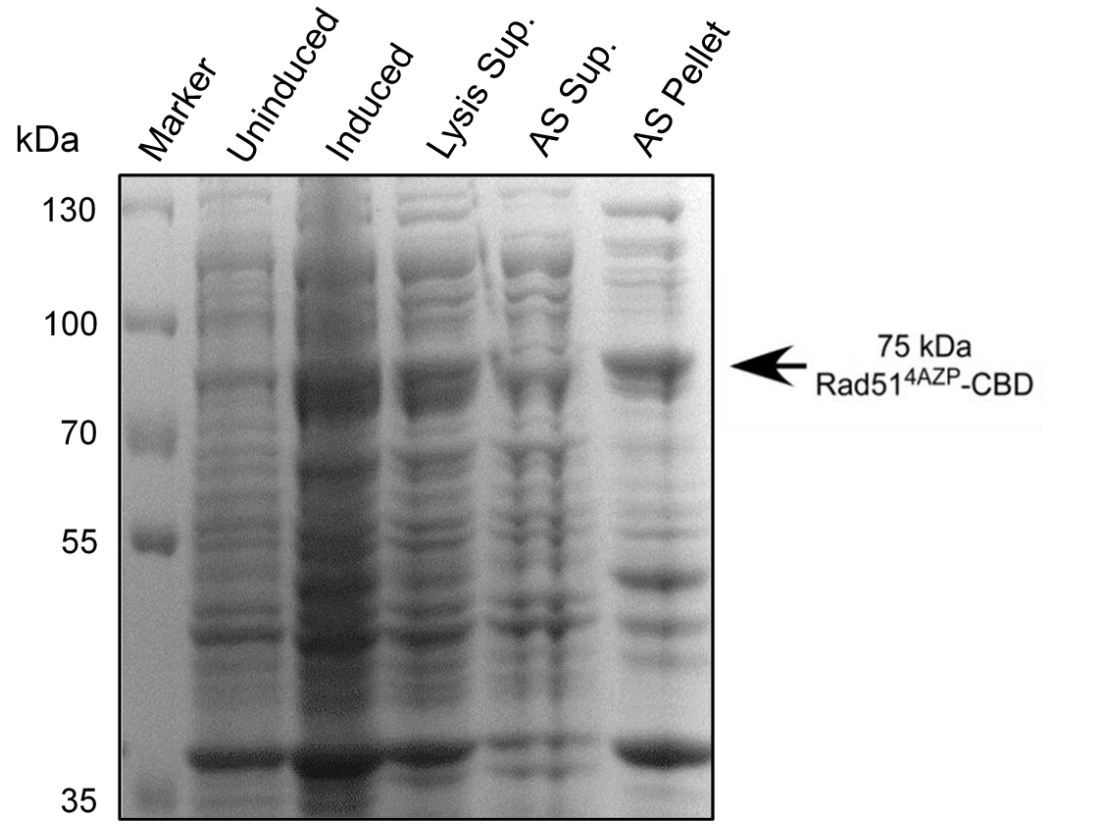



**Figure 8 | Ammonium sulfate precipitation of Rad51.** Upon the gradual addition of 0.24g/mL ammonium sulfate to the clarified lysate (Lysis Sup.), Rad51 is salted out of solution as a precipitate (AS Pellet), while multiple undesired proteins remain in solution (AS Sup.).

mL of resuspended protein in a 50 mL falcon tube, and shaken at 4ºC for 3-5 hours.

Next the beads were spun down at 4,000 RPM for 5 minutes. After careful removal of the supernatant, the beads were washed with AS Resuspension Buffer four times. Finally, 30 mM Dithiothreitol (DTT) AS Resuspension Buffer was prepared, 5 mL of which was combined with the beads and shaken at 4ºC overnight.

The following morning, the beads were again spun at 4,000 RPM for 5 minutes. The supernatant, containing the purified protein, was removed. Further 24 hour reactions with 30 mM DTT AS Resuspension Buffer were performed for a total of five elutions. The chitin elutions were pooled and concentrated using a 10 kDa concentrator (Pall Life Sciences).

It is important to note the significance of the chitin binding domain in purifying the probe. Purifying out the Rad51UAA from the truncated Rad51 product (where the amber stop codon was recognized by the release factor as an amber stop codon) presents a concern, as both proteins can behave similarly through purification steps. However, the truncated product will not have the C-terminal chitin binding domain, as translation stops prior to the coding for this domain. Hence, no truncated Rad51 protein product should bind to a chitin column. This technique works to separate out the Rad51 products (Fig. 9). After cleaving the CBD, the purified Rad51UAA can be eluted.

*2.3.4 Fluorescent Labeling*

Once concentrated, the protein, in buffer, was mixed with 10 molar excess of either Click-IT® Alexa Fluor® 555 DIBO Alkyne Copper-free fluorophore (Molecular Probes) or DBCO-Cy3 for Copper-free Click Chemistry (Sigma), both of which had been resuspended in Dimethylsulfoxide (DMSO). Click chemistry reactions are selective, and proceed well in water[63]. The labeling reaction was covered in aluminum foil, and rocked at 4ºC for 8 hours.

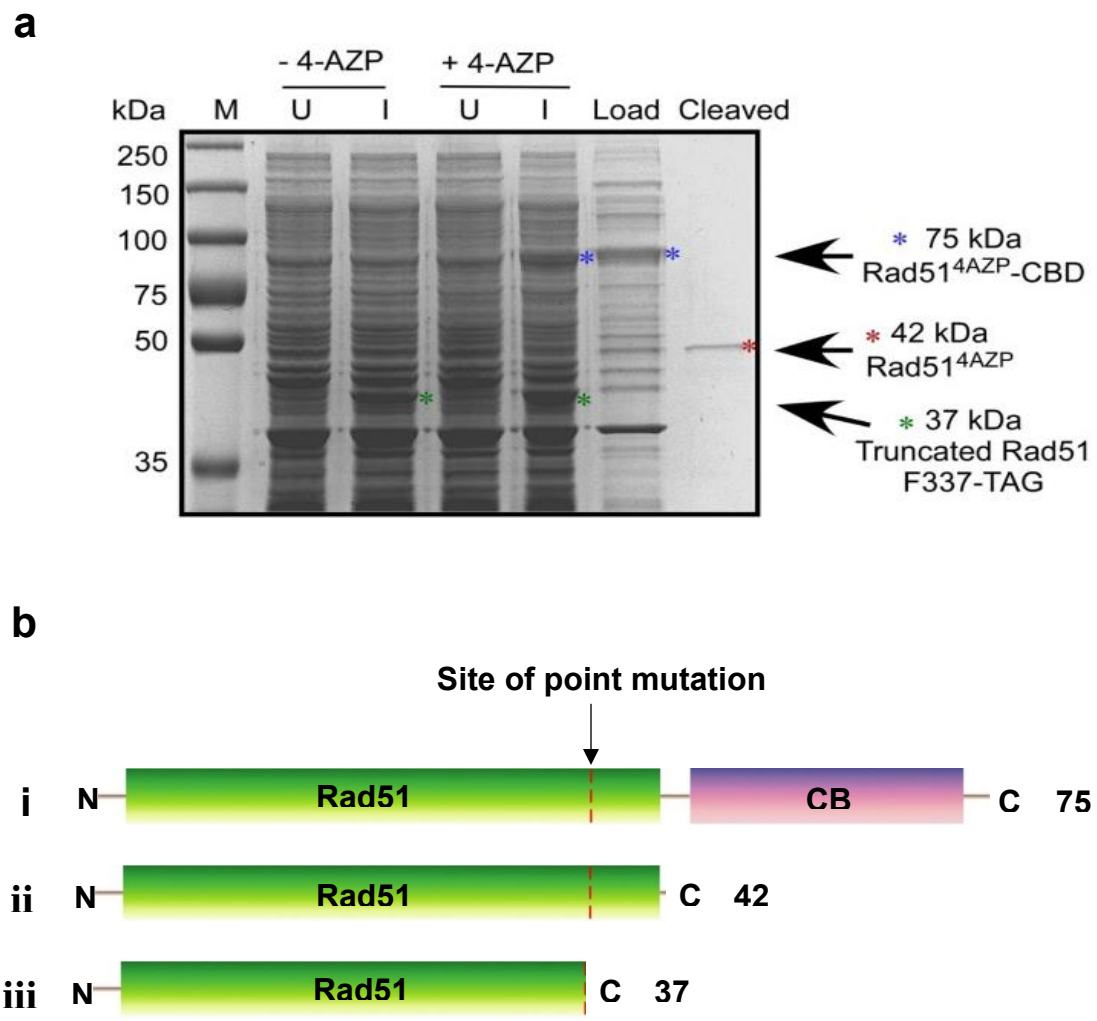



**Figure 9 | Rad51 purification with a chitin binding domain. a,** An 8% SDS page gel displaying the expression and chitin bead purification of Rad51. The first lane is marker. The next two lanes show protein expression in uninduced and induced cells, where there is no introduction of 4AZP. The two lanes following display, respectively, the same conditions, except 4AZP is introduced at induction. The last two lanes include samples loaded onto, and cleaved from, a chitin bead column. **b,** A schematic representation of the Rad51 protein. Every Rad51UAA construct has undergone a point mutation to code for a stop codon at the insertion point for the UAA. **i,** When the Rad51 successfully has the UAA incorporated, translation will not stop until after the inclusion of a chitin binding domain. This is seen in the above gel at 75 kDa. **ii,** After cleavage by DTT, the Rad51UAA (42 kDa) elutes from the Chitin column. **iii,** When the stop codon is read by the cell as a stop codon, not as a site for UAA incorporation, a truncated Rad51 protein is produced (37 kDa). This protein is unable to bind to a Chitin column.

Labeled Rad51 protein was next ran over a P6 Bio-Gel size exclusion column (Bio-Rad). Rad51 Size-Exclusion Buffer (20 mM Tris-HCl pH 7.5, 10% glycerol, 0.5 mM EDTA, 1 mM 2-mercaptoethanol (BME), and 100 mM NaCl) was used to equilibrate and run the column. After labeled protein entered the gel by gravity, Rad51 Size-Exclusion Buffer was ran through the column at a rate of 0.3 mL/minute. The appropriate fractions, as indicated by an 8% SDS-PAGE gel, were pooled and concentrated in a 10 kDa concentrator (Pall Life Sciences) to provide purified Rad51-Alx555 or Cy3.

*2.3.5 Synthesis of 4-Azido-L-phenylalanine*

A common barrier to the use of unnatural amino acids is their high cost and low availability[64]. As the purchase of 4-Azido-L-phenylalanine (4AZP) can be expensive at low quantities, an attempt was made to synthesize this unnatural amino acid from the precursor 4-amino-L-phenylalanine (4AMP). 4AMP is available in bulk, and at a substantially lower cost. Having a sufficient supply of 4AZP is critical in developing Rad51-UAA probes. During the expression of a Rad51-UAA probe, the synthetic tRNA, being responsible for the incorporation of 4AZP into the polypeptide chain, could read any amber stop codon in the mRNA undergoing translation as a site to insert 4AZP. Therefore, 4AZP must be supplied in excess during the expression to allow for it to be incorporated into a significant amount of Rad51.

The following method is specifically for 4AZP synthesis from 4AMP with a FMOC (Fluorenylmethyloxycarbonyl) blocking group (Fig. 10) attached to the

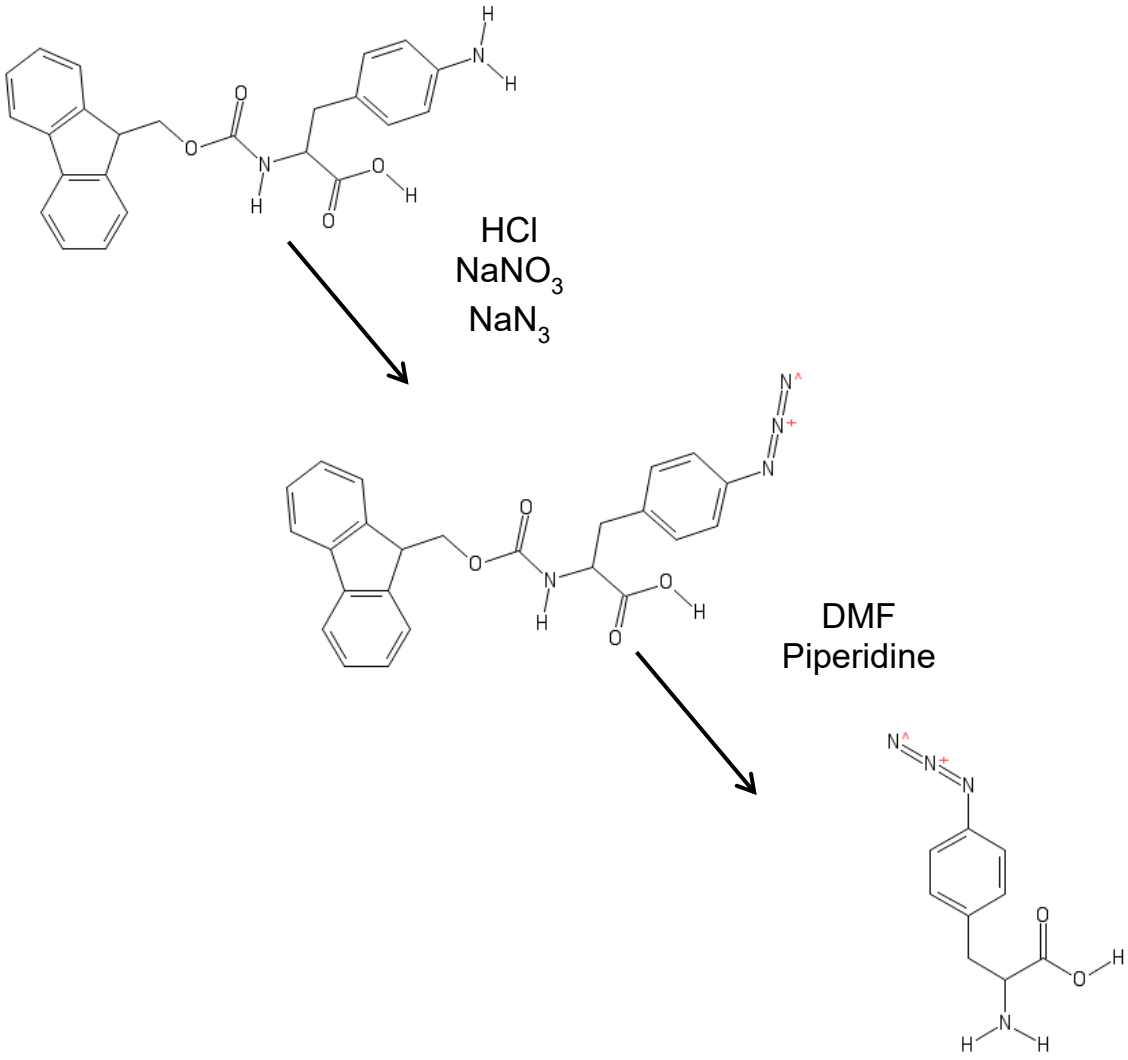



**Figure 10 | Synthesis of 4-azido-l-phenylalanine.** FMOC-4-amino-l-phenylalanine is converted into the intermediate, (2-((((9H-fluroen-9-yl)methoxy)carbonyl)amino)-3-(4-Azidophenyl)propanic acid where the side chain amino group has been altered to an azide, by the acidic addition of nitrate and azide. This stable intermediate is further synthesized to 4-azido-l-phenylalanine by the removal of the FMOC blocking group by piperidine dissolved in DMF. Images were generated at *pubchem.ncbi.nlm.nih.gov*

amino group of the unnatural amino acid. A blocking agent is necessary to ensure that, during the first step of synthesis, an unwanted reaction does not proceed with the amino acid's backbone amino group. The Dr. Tom Chang Lab at Utah State University assisted in developing, and carrying out, this protocol. Synthesis was performed with 5 grams of starting material, as larger-scale reactions gave smaller yields. All reactions were mixed and carried out in a fume hood.

With 5 g of starting material, 50 mL of $H_20$ was used as a solvent to give a 4AMP concentration of 12.42 mM. Much of the starting material remained solid. After the addition of 11.2 mL concentrated HCL (37% volume per volume), the starting material dissolved completely into solution. This was followed by cooling to 4˚C and the addition of $NaNO_2$ to a concentration of 12.42 mM (giving a 1 : 1 ratio with the 4AMP), which was first dissolved in 10 mL water. Following 10 minutes of stirring, $NaN_3$ (14.49 mmol, providing a 1.2 : 1 ratio with 4AMP) was added, slowly and cautiously, over a period of 10 minutes, while still stirring. The solution was stirred in the fume hood, overnight, at room temperature. On occasion, gradual addition of $NaN_3$ resulted in the formation of a foam, preventing proper mixing of the compounds. When this occurred, a small volume of ethyl acetate (approximately 5 mL) was added to the solution, bringing the foam back into solution and allowing for proper mixing.

Azides, whether organic or inorganic, can be very sensitive and explosive with minimal energy input. Therefore, the addition of an azide must be performed under extreme caution.

The following morning, a white precipitate could be observed. Next, the solution underwent three 100 mL ethyl acetate extractions. With each extraction, the ethyl acetate was mixed with the reaction in a 500mL seperatory funnel. The solution was agitated by manual shaking, with frequent breaks to alleviate the positive pressure built inside the funnel. The extractions were combined and washed with 50 mL of $H_20$ twice, and once with 50 mL of brine solution. The organic layer was collected and dried with anhydrous $MgSO_4$ (to remove any left-over $H_2O$), although $Na_2SO_4$ could also be used. Following a filtration to separate out the drying-salt, the removal of ethyl acetate was performed with the use of a rotovap apparatus. The product formed was a white powder that was allowed to dry overnight (under negative pressure). The resulting product was dry (2-((((9*H*-fluroen-9-yl)methoxy)carbonyl)amino)-3-(4-Azidophenyl)propanic acid (Fig. 10). Weighing the dried intermediate gave a product yield of 76.2%.

Following the drying of this intermediate product, it was dissolved in 9 mL of DMF (dimethylformamide). After the compound was completely dissolved, 9 mL of piperidine, in 27 mL DMF, was added. The solution was allowed to stir overnight at room temperature. The next morning a white-beige precipitate of the final product was observed. This precipitate was collected using a filtered Hirsh funnel, and afterwards washed with 25 mL of DMF. The resulting purified product

was white, and dried on the same funnel under vacuum. After 2-3 days the product was removed from the filter paper by scraping with a spatula. The final yield of the synthesized 4AZP was 52%. A C-13 NMR of the final product is shown in Supplementary Figure 1.

*2.3.6 Stopped flow Instrumentation*

Stopped flow instrumentation was utilized to obtain real time data for the probes. The basic setup of a Stopped flow device has two syringes containing sample solutions (Fig. 11). A drive force combines the two samples in a mixer compartment, which quickly arrive at the cell. A light source penetrates through the cell, and absorbance can be measured at an angle of 180 degrees, while fluorescence can be detected at 90 degrees. After the period of data collection for the reaction has subsided, the sample is discarded from the cell. A stopping syringe helps control the volume that is mixed from the two syringes.

## 2.4 Results and Discussion

*2.4.1 Labeled Probes*

The first probe developed was the Rad51 337-UAA. The successful fluorescent labeling of this probe, completed through click chemistry with an alkyne Alx555, is shown in Figure 12. Coomassie stain and fluorescent images of the same gel show that the Alx555 was indeed incorporated. Next, the additional Rad51 residues F290, S291, and Y16 were successfully mutated, expressed, purified, and labeled with Alx555 alkyne dye (Fig. 13-15).

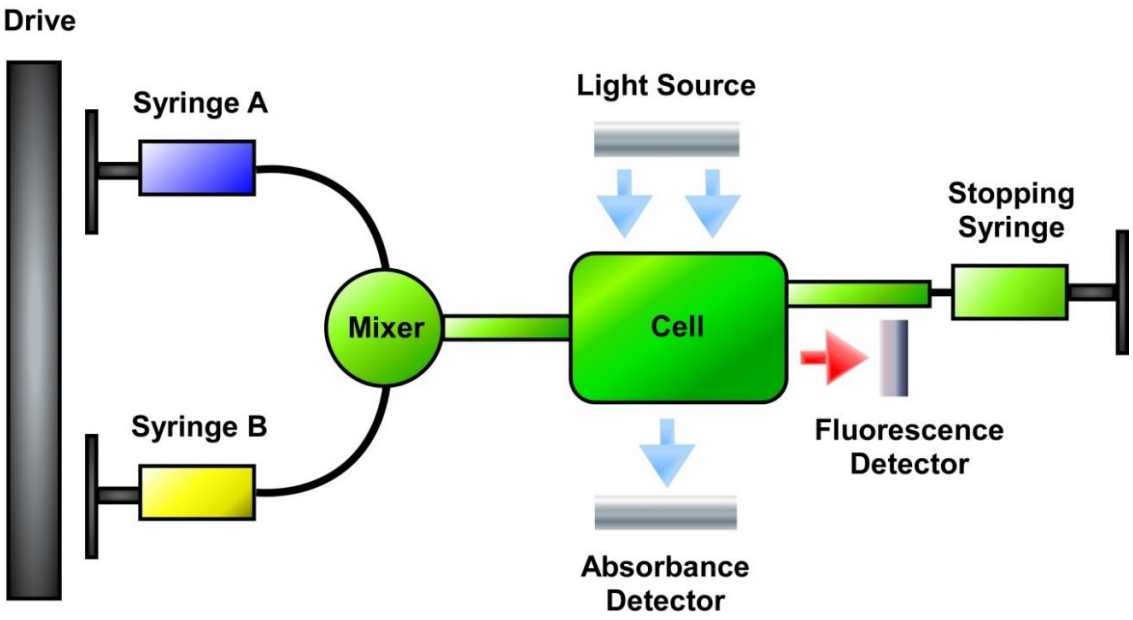



**Figure 11 | An overview of a Stopped flow experimental setup.** Two samples are loaded, one into syringe A and one into syringe B. A drive forces equal amounts from both syringes, into the mixer and then to the cell where a light source, at a given wavelength, penetrates the mixed solution. An absorbance detector is set up at a 180º angle from the incoming light, while a fluorescence detector is set at 90º. A stopping syringe helps control the volume mixed in each run.

*2.4.2 Rad51-UAA Stopped flow Experiments*

A Stopped flow experiment can yield helpful data, as the two solutions may contain interactive species (such as Rad51 in one well and DNA in the other) that undergo an activity upon mixing. This technique is specifically useful when an activity occurs that results in a change in the absorbent or fluorescent characteristics of the two samples. A benefit of employing Stopped flow technique is that the data can be measured nearly instantaneously from mixing.

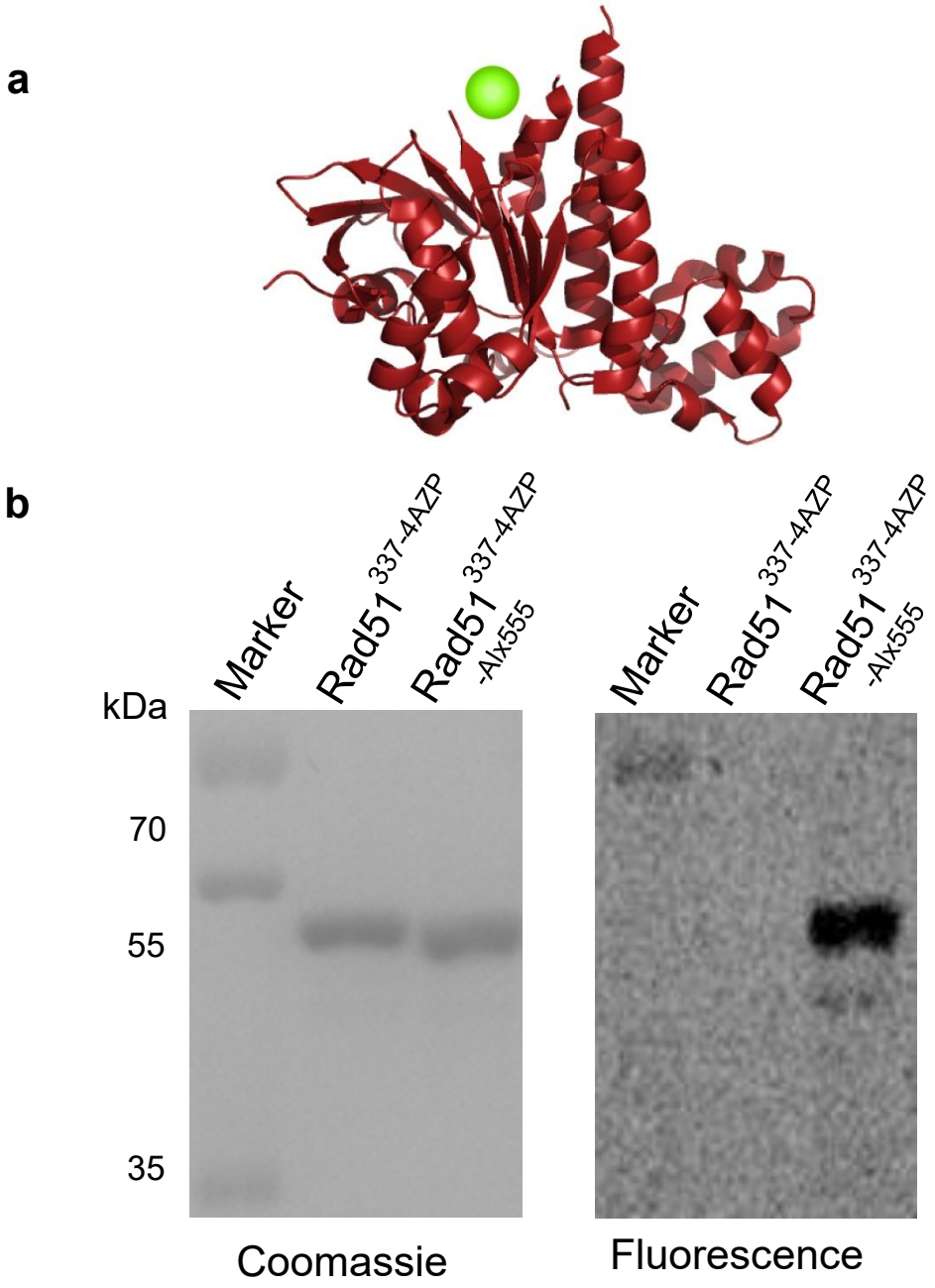



**Figure 12 | Fluorescent labeling of Rad51-337-4AZP. a,** The placement of the probe at the disordered loop L2. **b,** Both Coomassie and fluorescent images were taken of the same 8% SDS PAGE gel. Rad51-337-4AZP is viewable in the Coomassie stain image, regardless of whether or not a fluorophore has been added. The fluorescent image indicated that the dye (Alx555) has successfully been linked to the probe. Structure was generated from PDB 3LDA.

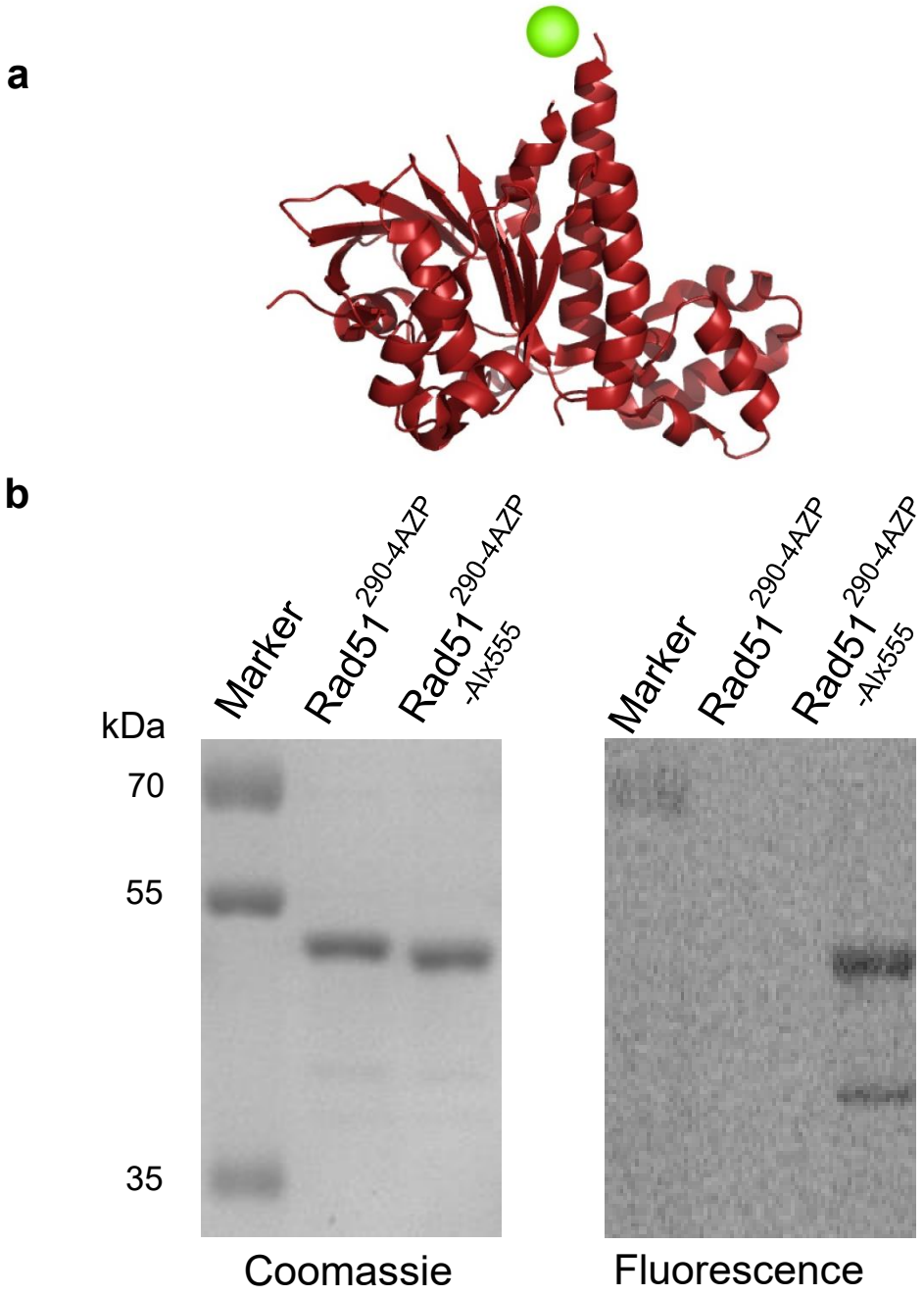



**Figure 13 | Fluorescent labeling of Rad51-290-4AZP. a,** The probe is placed at the disordered loop L1. **b,** Coomassie and fluorescent images of the same 8% SDS PAGE gel. Rad51-290-4AZP is seen in the Coomassie image, regardless of fluorophore incorporation. The fluorescent image demonstrates that the probe has been successfully labeled with the fluorophore (Alx555). Structure was generated from PDB 3LDA.

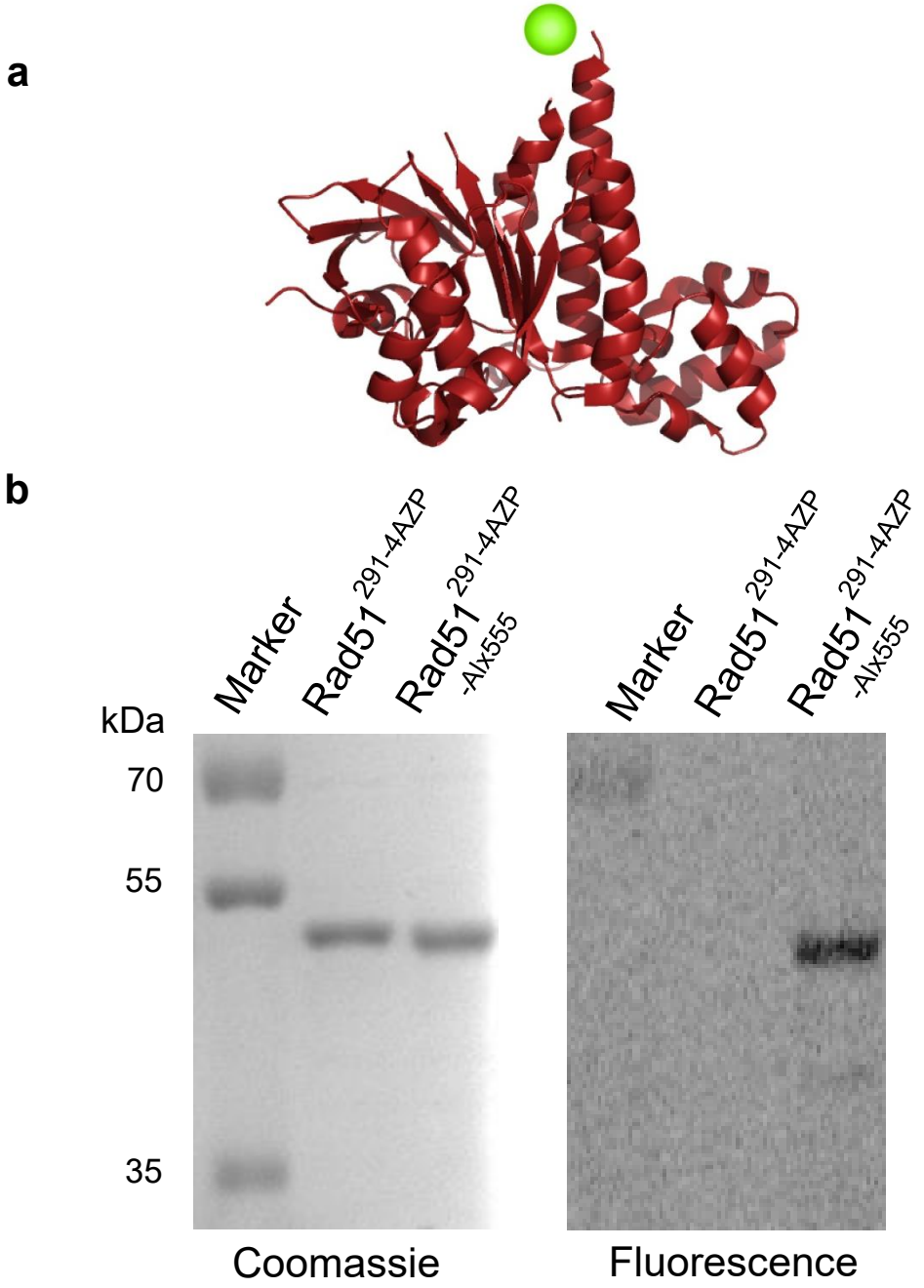



**Figure 14 | Fluorescent labeling of Rad51-291-4AZP. a,** The placement of the probe at the disordered loop L1. **b,** Coomassie and fluorescent images of the same 8% SDS PAGE gel. Rad51-291-4AZP is seen with Coomassie stain, regardless of whether or not a fluorophore has been added. Successful addition of Alx555 is seen in the fluorescence image. Structure was generated from PDB 3LDA.

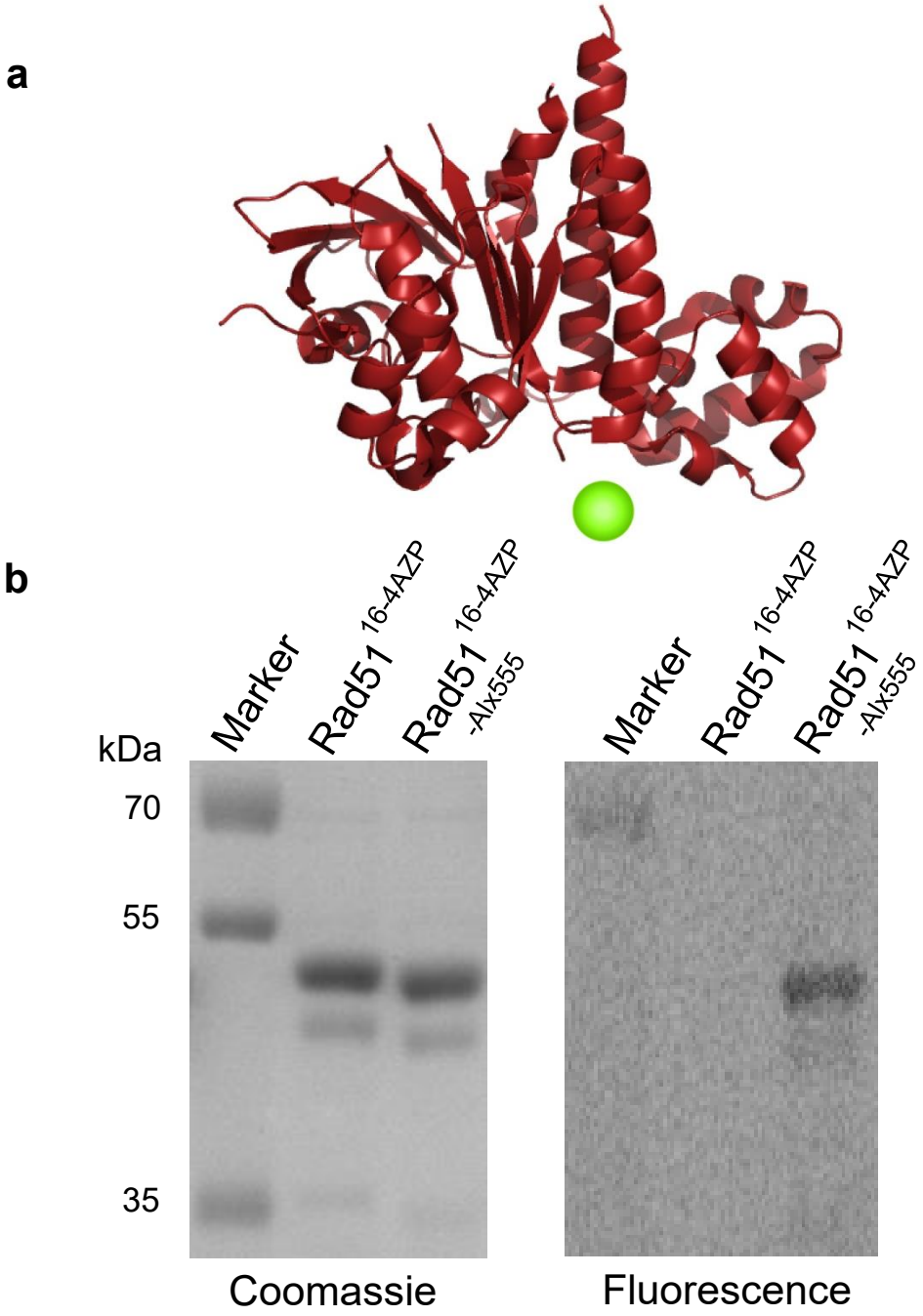



**Figure 15 | Fluorescent labeling of Rad51-16-4AZP. a,** The placement of the probe at N terminus. **b,** Coomassie and fluorescent images of the same 8% SDS PAGE gel. Rad51-16-4AZP is seen with Coomassie stain, regardless of whether or not a fluorophore has been added. The fluorescence image indicates that the dye (Alx555) has been successfully incorporated. Structure was generated from PDB 3LDA.

This can provide the kinetic parameters of an interaction, as well as provide insights into what could be occurring in the reaction mechanism.

The first Stopped flow experiment utilizing the Rad51 probe was performed to compare the binding behavior of the Rad51-337-UAA with the wild type version of the protein (Fig. 16a). Experiments were performed with Stopped flow instrumentation, as displayed in Figure 11. Syringe A held $dT_{79}$ ssDNA with a Cy3 fluorescent label on the 5′ end, ATP, and reaction buffer (50 mM Tris-HCl pH 7.5, 50 mM NaCl, and 10 mM $MgCl_2$). Syringe B contained wild type Rad51 and buffer. Upon mixing, a change in the Cy3 fluorescence was observed. A trial of the experiment followed with newly synthesized Rad51UAA replacing the wild type Rad51. As observed in the data (Fig. 16a), the Rad51UAA shows similar fluorescence quenching behavior as the wild type. This indicates that the incorporation of 4AZP does not alter Rad51 filament formation.

Next, a second Stopped flow experiment included Rad51-337-Alx555 and buffer in one syringe, with ATP and buffer in the other (Fig. 16b). Fluorescence signal strength was recorded upon mixing. To measure the change in fluorescence emitted by the probe upon filamentous activity, the same conditions were repeated, save with dT70 DNA included in the second syringe. As displayed in the figure, a change in signal strength is observed when DNA is available. This intensity change comes about as the Rad51 filament forms on DNA. Because it exhibits a change in signal upon filament formation, the Rad51UAA could potentially act as a probe for when HR is being promoted (and

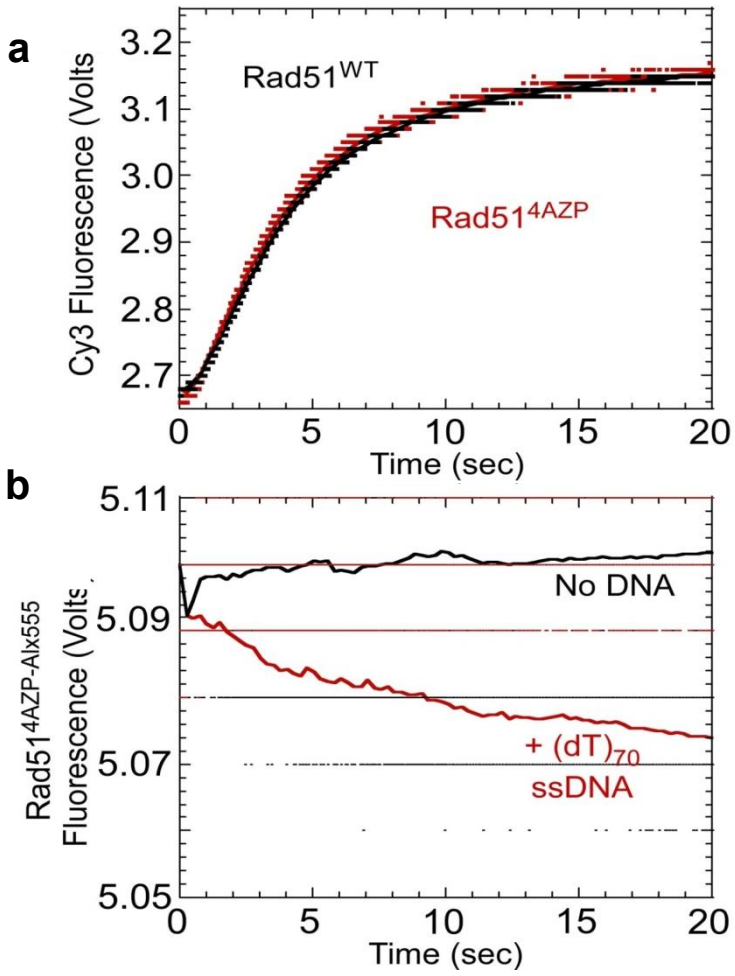



**Figure 16 | Rad51-337-4AZP DNA binding. a,** Displayed are identical binding kinetics for both wild type Rad51, and Rad51 with 4AZP incorporated. **b,** The fluorescence of the Rad51-4AZP-Alx555 protein changes upon binding to single stranded $dT_{70}$ ssDNA.

hence undergoing filamentous activity) or antagonized (and the filament is being disassembled) by mediator proteins. This experiment was repeated with Rad51-290-Cy3, Rad51-291-Cy3, and Rad51-337-Cy3 (Fig. 17). As the seen in the

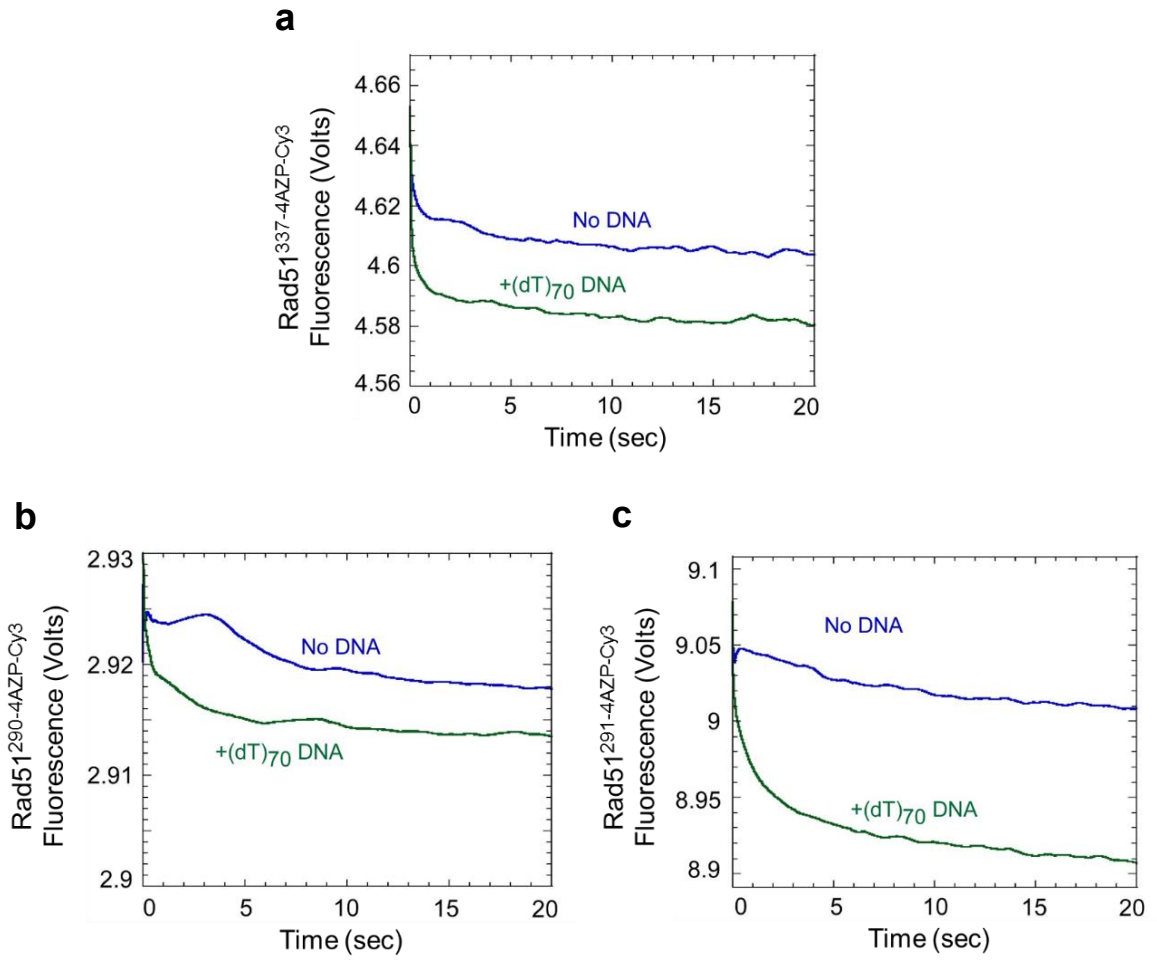



**Figure 17 | Rad51-4AZP-Cy3 Stopped flow experiments.** These experiments indicate a change in fluorescence emission intensity for the probes **a** Rad51-337-Cy3, **b** Rad51-290-Cy3, and **c** Rad51-291-Cy3 upon binding to $dT_{70}$ DNA.

figure, a change in fluorescence emission intensity is observed when Stopped flow reactions include DNA. These preliminary experiments indicate that the three probes are providing a change in signal upon DNA binding, allowing for the visualization of Rad51-DNA filament formation.

## 2.5 Future Work

All Rad51 probes will have their activity levels measured in comparison to wild type protein. This will be performed by gel mobility shift assays where the Rad51 protein catalyzes a strand exchange reaction between plasmid DNA with a shorter, complimentary sequence of phosphorus-32 labeled DNA. To be accurate reporters of HR filament dynamics, Rad51-4AZP-fluorophore probes must demonstrate activity at the same level as the wild type. This experiment is further explained in Chapter 4.

The 16-4AZP probe will be labeled with two separate fluorophores, Cy3 and Cy5, and used for FRET experimentation. By having equal amounts of each fluorophore-probe, FRET interactions will exist upon filament formation and some of the probe will exhibit a signal. When Rad51 monomers are not interacting, the distance between Cy3 and Cy5 will be too great in relation to Foster radius (54-61 Angstroms[53]) for FRET to occur. This experiment will be performed with the use of a fluorometer. See Figure 18 for a model of the proposed experiment. Initially, Rad51-16-Cy3 and Rad51-16-Cy5 will be incubated with ATP, all in excess. A baseline of FRET signal will be taken, after which ssDNA will be introduced to the system. Upon addition of the oligonucleotide, FRET signal should increase as many of the individual proteins making up the filament will have an adjacent protein with a different fluorophore which is compatible for FRET. Once all available ssDNA has been bound, a plateau should be seen in the FRET signal. The same experiment done with ssDNA of various lengths will

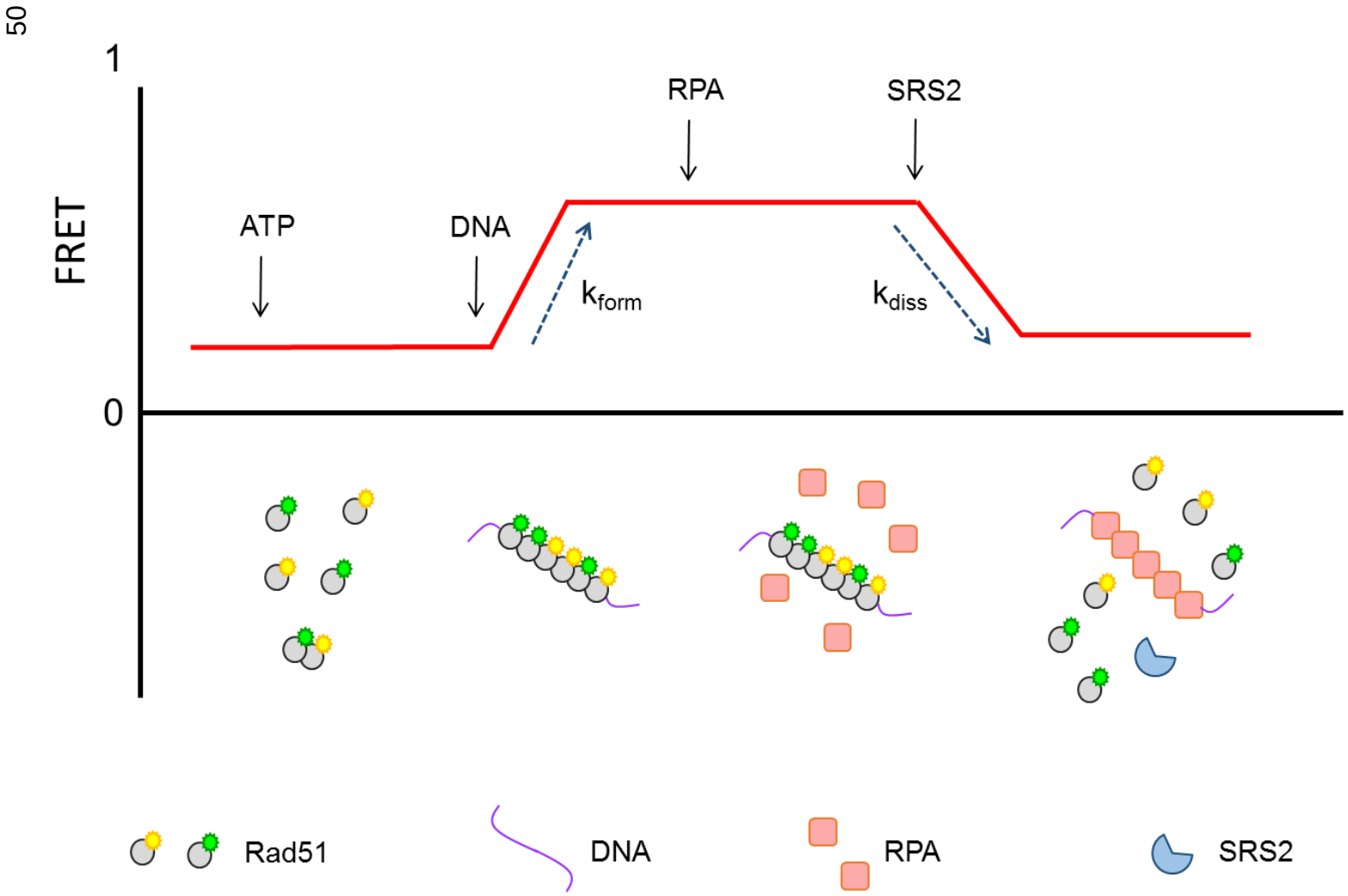



**Figure 18 | An overview of a proposed FRET experiment.** As DNA is mixed with ATP, along with Rad51-Cy3 and Rad51-Cy5, an increase in FRET signal should be seen as the Rad51-DNA filament forms. After the signal is saturated, as all DNA is bound, the introduction of SRS2 will disassembly the filament and RPA will be allowed to bind the free DNA.

provide rates of filament formation as a function of length (in base pairs), and the rate of formation (in base pairs $s^{-1}$) can be extracted.

Next, RPA will be introduced. Initially, no change should be observed in the FRET signal. Finally, SRS2, an anti-recombinase, is added. SRS2 works to dissociate the Rad51 filament[37], allowing RPA to bind the free ssDNA. Accordingly, the FRET signal should then fall back to baseline levels once SRS2 has fully removed all Rad51 filaments. Similarly, to the addition of ssDNA, the slope of the data as the signal falls, from the point of SRS2 introduction to reaching near-baseline levels, provides a rate for filament dissociation.

# CHAPTER III

# A SECONDARY FLUORESCENT PROBE WITH PLASMODIUM FALCIPARUM SINGLE STRAND DNA BINDING PROTEIN

## 3.1 Abstract

*DNA's integrity is vital to the health of the cell, and must be heavily protected. At instances where DNA is single stranded (ssDNA), it becomes more vulnerable to damage as it has lower stability. Single strand DNA binding protein (SSB) works to protect single stranded DNA from degradation and damage. Plasmodium falciparum SSB forms a tetramer in solution and demonstrates only one binding mode. In this work, PfSSB is converted into a fluorescent probe for unbound ssDNA by labeling it with a maleimide-coumarin fluorophore. Attaching a fluorophore to PfSSB results in a probe that quickly associates with ssDNA, and functions to monitor both the helicase and anti-recombinase function of the homologous recombination protein, SRS2.*

## 3.2 Introduction

The single strand DNA binding protein is necessary in every domain[65]. As the name suggests, this protein binds to free ssDNA in the cell, doing so independent of sequence[66]. SSB works to protect DNA from being degraded. This protein provides a significant function by binding ssDNA during DNA metabolism. SSB makes for a positive candidate to be converted into a probe which quantifies when free ssDNA is available in the cell, as it is both commonly found and binds with very high affinity to ssDNA.

*E. coli* SSB forms a tetramer[67,68], and has been utilized as a probe for helicase activity by binding free ssDNA that has been unwound by said helicase[69]. Upon binding to ssDNA, the level of intrinsic tryptophan fluorescence of the SSB tetramer is altered. The fluorescence change occurs as the ssDNA alters the immediate microenvironment of the exterior tryptophans on each subunit.

While intrinsic tryptophan fluorescence can be a useful technique to measure SSB-ssDNA binding, the probe was improved upon by the addition of a fluorophore[70]. This labeled probe undergoes a greater change in fluorescence upon binding to ssDNA than the exterior tryptophans, producing greater sensitivity in experiments.

Interestingly, the *E. coli* SSB tetramer does not always bind the same number of nucleotides[71]. Depending on the NaCl concentration, two binding states have been identified, one where the tetramer associates with 65 nucleotides at high NaCl concentration, and another where the SSB associates with 35 nucleotides at lower NaCl concentrations[72]. A similar probe has recently been developed with the species *Bacillus subtilis*[73]. Both of these SSB probes share a similarity in that they exhibit differing binding modes dependent upon salt concentration.

The differences between the two binding modes create an issue in experiments. The buffer conditions must be regimented. Also, if multiple proteins are present in a reaction with the probe, it must first be ruled out that these

proteins are not affecting the binding mode of SSB. Both of the aforementioned probes rely on mutations to introduce novel cysteines to the protein. Changing the amino acid residue at a given sight is fundamentally invasive, and can have unforeseen effects on the protein's activity.

SSB makes for a good candidate to study HR dynamics for multiple reasons. Because it binds DNA with a high reported affinity ($10^{10}$ $M^{-1}$)[72], it can act as a reporter for when a protein, such as Rad51 or RPA, is removed from ssDNA (as long as it is not competing either protein off of the ssDNA). As pro- and anti-recombinases can act by directly promoting or removing the Rad51 filament, SSB could act as a second messenger, indicating which proteins are removing Rad51 from ssDNA. This can be done by introducing SSB and proteins thought to be anti- recombinases to a preformed Rad51-ssDNA filament. If the SSB probe demonstrates a change in fluorescence, this would indicate that it has bound to ssDNA, and hence that the presumed anti- recombinase works to directly dissociate the Rad51-DNA filament. Alternatively, any change in fluorescent signal could be brought about by protein-protein interactions involving SSB. Therefore, proper controls should be
performed to ensure that such an interaction is not taking place between SSB and the other proteins involved in the experiment.

Additionally, helicase activity can be monitored by an SSB probe. Helicase activity is important in HR. Helicases are used in the resecting of the two broken ends of the damaged double stranded DNA (dsDNA), as well as in strand

invasion of the ssDNA overhangs. An SSB probe could function to monitor helicase activity, as SSB binds to the freshly unwound ssDNA.

*Plasmodium falciparum* is a eukaryotic parasite, and can cause malaria in humans. This organism also creates its own version of the single-stranded DNA binding protein, *Pf*SSB. This eukaryotic SSB shares similarity to *Ec*SSB, in that it forms a tetramer in solution[74]. Likewise, the manner in which ssDNA wraps around *Pf*SSB can be related to the seams along the outside of a tennis ball (Fig. 19). The protein has a binding core responsible for the ssDNA binding behavior. Significantly, *Pf*SSB does not display salt-dependent binding modes[75]. *Pf*SSB binds around 65 nucleotides, regardless of salt concentration. This allows for data to be gathered reliably across a wide range of salt concentrations. *Plasmodium falciparum* SSB has one cysteine residue on each subunit of the tetramer, residing at its DNA binding core, making it a good candidate for fluorescent labeling with a maleimide fluorophore. Maleimide fluorescent labeling was discussed in Chapter 2. In the solved crystal structure of *Pf*SSB[74] (Fig. 19), this cysteine residue is near the ssDNA-binding region and on the protein's surface. Maleimide chemistry is fast and specific for thiols. Unbound, surface exposed cysteine residues contain a sulfhydryl group that reacts with the carbon-carbon double bound of the maleimide group to directly link the side chain with a maleimide fluorophore.

With one solvent exposed cysteine residue, and the display of only one type of ssDNA binding mode, *Pf*SSB makes a good candidate for a novel probe

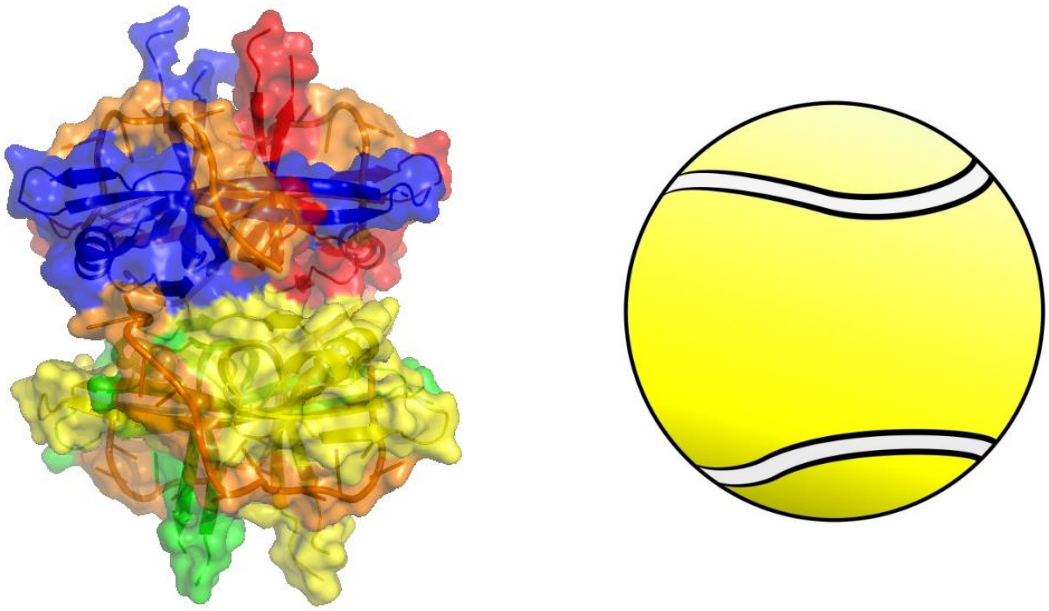


**Figure 19 | *Pf*SSB binding to ssDNA.** Observed to the left is the crystal structure for the *Pf*SSB tetramer (PDB, 3ULP), with subunits in blue, red, yellow, and green, bound with two short strands of $dT_{35}$ ssDNA (orange). The manner in which ssDNA tightly wraps around *Pf*SSB is compared to the seams along the surface of a tennis ball.

in studying HR dynamics. Because *Pf*SSB binds to ssDNA at a similar ratio regardless of NaCl concentration, the amount of NaCl in a given experiment can be adapted for the specific proteins involved, allowing for greater flexibility in potential experiments.

## 3.3 Materials and Methods

### *3.3.1 PfSSB Purification*

The sequence for the *Pf*SSB gene was previously incorporated into a pET21 vector. The pET21 vector contains ampicillin resistance as well as a t7 promoter for inducible protein expression. The vector was transformed into Rosetta™(DE3)pLysS Competent Cells (Novagen) for protein expression. The

cells were grown up in 1L Luria Broth (LB) with 100 µg/mL ampicillin, shaken at 37ºC and 250 RPM. Purification steps can be seen in Figure 20. At an optical density, measured by a spectrophotometer at 600 nm ($OD_{600}$), of 0.6, transcription was induced by the addition of 400 µM Isopropyl β-D-1-thiogalactopyranoside (IPTG). After three additional hours of shaking, the growth was spun down at 4,000 RPM for 30 minutes and the cell pellet stored at -80ºC. Cell pellets containing the *Pf*SSB protein were re-suspended in 200 mL of lysis buffer (10% sucrose, 50 mM Tris pH 8.3, 1 mM EDTA, 0.2 M NaCl, and 15 mM Spermidine), along with 100 µM phenylmethanesulfonyl fluoride (PMSF) dissolved in isopropanol. The resuspension was transferred to a beaker, and mixed on a stir plate at 4ºC for 10 minutes. Afterward, fresh lysozyme (Hen egg white) was added to a final concentration of 20 mg/100 mL. PMSF was added a second time, to a final concentration of 200 µM. The solution was then stirred at 4ºC for another 30 minutes. Next, 6 mL of 100 mM sodium deoxycholate, dissolved in 0.125 M NaOH, was added. This was followed by 20 minutes of stirring at 4ºC.

Sonication followed to thoroughly lyse the *E. coli* cells. The sonicator was set to 50% duty cycle. Cells were sonicated for one minute, followed by one minute of rest. This was repeated twice for a total of three minutes of sonication. The resulting cell lysate was centrifuged at 13,000 RPM for 80 minutes.

A Polimin P precipitation was performed with the supernatant. A Polimin P solution [5% volume/volume Polimin P HCL, pH 8.3 and 10% volume/volume of

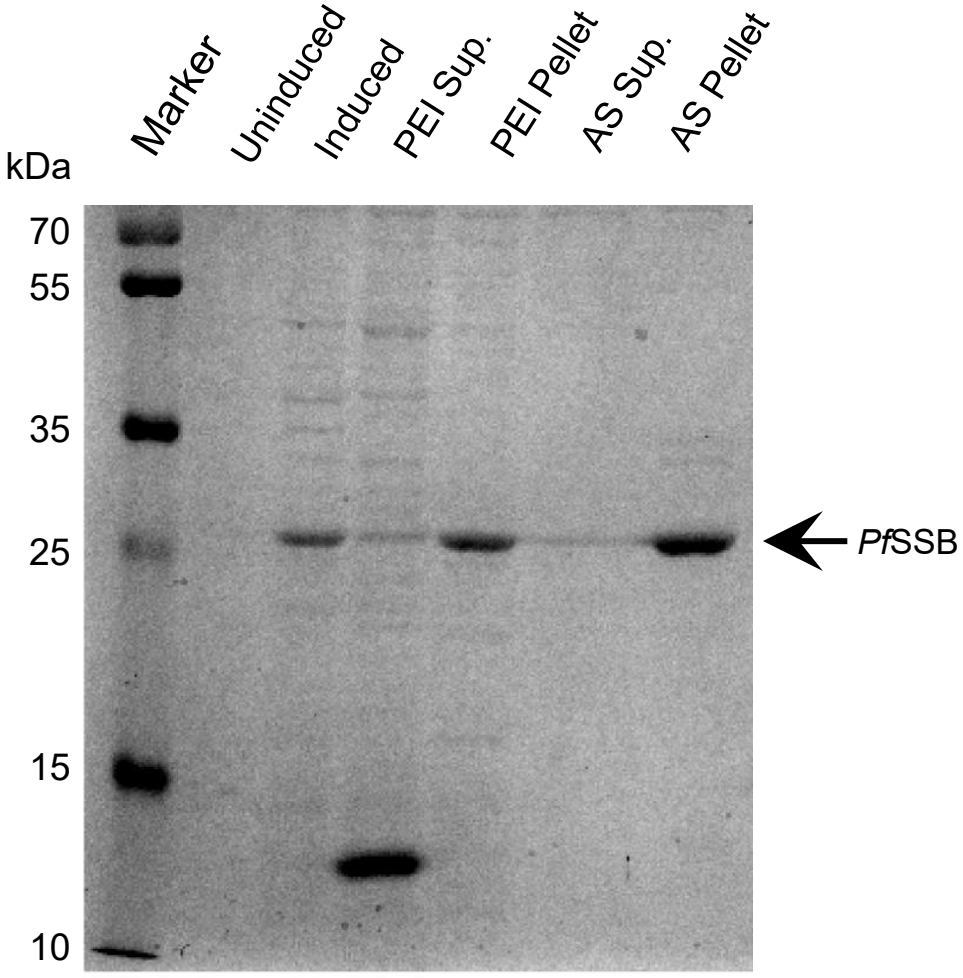



**Figure 20 | *Pf*SSB purification gel.** A 12% SDS PAGE gel was ran, including uninduced and induced protein expression, as well as the supernatant and pellet from each purification step. Purified *Pf*SSB can be viewed in the final, ammonium sulfate pellet lane.

50% weight/volume Polyethylenimine (PEI)] was added, slowly, to a final concentration of 0.2% Polimin P, while being continuously stirred at 4ºC. After 15 minutes, the solution was centrifuged for 30 minutes at 10,000 RPM. The pellet was then resuspended in 225 mL of 0.4 M NaCl Buffer T (5% glycerol, 50 mM

Tris pH 8.3, 1 mM EDTA, and 0.4 M NaCl), and stirred at 4ºC for 25 minutes. This resuspension was again centrifuged for 30 minutes at 10,000 RPM.

To the supernatant, 30.8% weight/volume of ground Ammonium Sulfate crystals were added. This was performed slowly, over a 15 minute period, while the solution was stirring at 4ºC. This concentration of ammonium sulfate forces the *Pf*SSB protein to be salted out of the solution so that, after stirring for an additional 30 minutes, the protein can be collected by centrifugation. After spinning for 30 minutes at 10,000 RPM, the Ammonium Sulfate pellet was resuspended in 200 mls of 0.3 M NaCl Buffer T (5% glycerol, 50 mM Tris pH 8.3, 1 mM EDTA, and 0.3 M NaCl). One additional centrifugation, again at 10,000 RPM for 30 minutes, was performed in order to remove insoluble impurities.

The next step in purifying of the *Pf*SSB involved the use of a 125 mL ssDNA column. The ssDNA column was equilibrated with 100 mL 0.3 M NaCl Buffer T, to ensure that the column would be under the same buffer conditions as the resuspended protein. The protein was then loaded at 1mL/min onto the column, which was washed with 100 mL 0.3 M NaCl Buffer T. Next, a salt gradient was ran through the column to elute the bound protein. The elution buffer changed linearly from 100% to 0% 0.3 M NaCl Buffer T, as well as 0% to 100% 2.0 M NaCl Buffer T, with a total volume of 200 mL. The column was ran at 1mL/min. A following 10% SDS PAGE gel indicated which elutions contained the SSB protein.

The protein was concentrated in a 10 kDa Macrosep Advance concentrator (Pall Corporation), down to 3.5 mL. After measuring the concentration, the solution was dialyzed into Labeling Buffer (10% glycerol, 0.5 M NaCl, 30.75 mM $K_2HPO_4$, and 19.25 mM $KH_2PO_4$) overnight. The protein was flash-frozen with liquid nitrogen and stored at -80ºC for future labeling.

*3.3.2 PfSSB Fluorescent Labeling*

5mg/mL TCEP [(Tris)2-carboxyethyl]phosphine hydrochloride] was prepared in labeling buffer, and then added to the protein in a 16-fold molar excess. TCEP is a reducing agent, and is capable of reducing disulfide bonds. This is necessary as the thiol containing side chains of cysteine residues must be unbound so they can react upon mixing with the maleimide fluorophore. The protein was labeled in a 10-fold molar excess with the specific fluorophore 7-Diethylamino-3-[N-(2-maleimidoethyl)carbamoyl]coumarin (MDCC) (Life Technologies). 1 mg of dye was suspended in 200 µL dimethyl sulfoxide (DMSO), giving 13.04 mM MDCC. 160 µL of suspended dye was then added to the protein, with a final concentration of 2.98 mM MDDC for the labeling reaction. The tube was then covered in aluminum foil and placed on a rocker set to 15 RPM at 4ºC overnight. 1.25 µL of 2-Mercaptoethanol (BME), resulting in a 5-fold molar excess of 14.9 mM, was added to quench the reaction the following morning. BME works to quench the reaction as the thiol group competes with the side chains of free cysteine residues to react with the maleimide group.

*Pf*SSB-MDCC was then separated on a P-4 Bio-Gel (Bio-Rad) size exclusion column with Labeling Buffer. Collected protein was spun down to 0.5 mL in a 10 kDa concentrator (Pall Corporation). Labeled protein was dialyzed over the weekend into SSB Storage Buffer (50% glycerol, 20 mM Tris, pH 8.3, 1 mM EDTA, 0.5 M NaCl, and 1 mM BME). Labeled protein can be observed in both a coomassie and fluorescence image of the same 10% SDS PAGE gel (Fig. 21). Figures 21-25 incorporate images previously published by the author in PLoS ONE.[76]

After dialysis, the volume of *Pf* SSB-MDCC was 1.7 mL. Measuring the concentration of the fluorophore portion of the flow-through revealed two absorbance peaks, one peak near 430 nm and one peak near 270 nm, with the ratio of the magnitude of absorbance at these two peaks being approximately 1: 0.3634 (430 nm peak : 270 nm peak). While measuring the concentration of protein, the absorbance at 270 nm exhibited by the fluorophore must be accounted for, as this adds to the absorbance of the protein peak near 270 nm. Therefore, when an apparent *Pf*SSB-MDCC concentration of 56 μM for the tetramer is observed, the actual concentration is calculated to be 35.69 μM.

### *3.3.3 Annealed DNA*

As the SRS2 protein prefers binding to 3′ single stranded overhangs[77], an oligonucleotide, consisting of 25 nucleotides of double stranded DNA with a 16 nucleotide 3′ single stranded overhang was prepared by annealing two partially complementary strands of ssDNA. One strand was 25 nucleotides in length, and

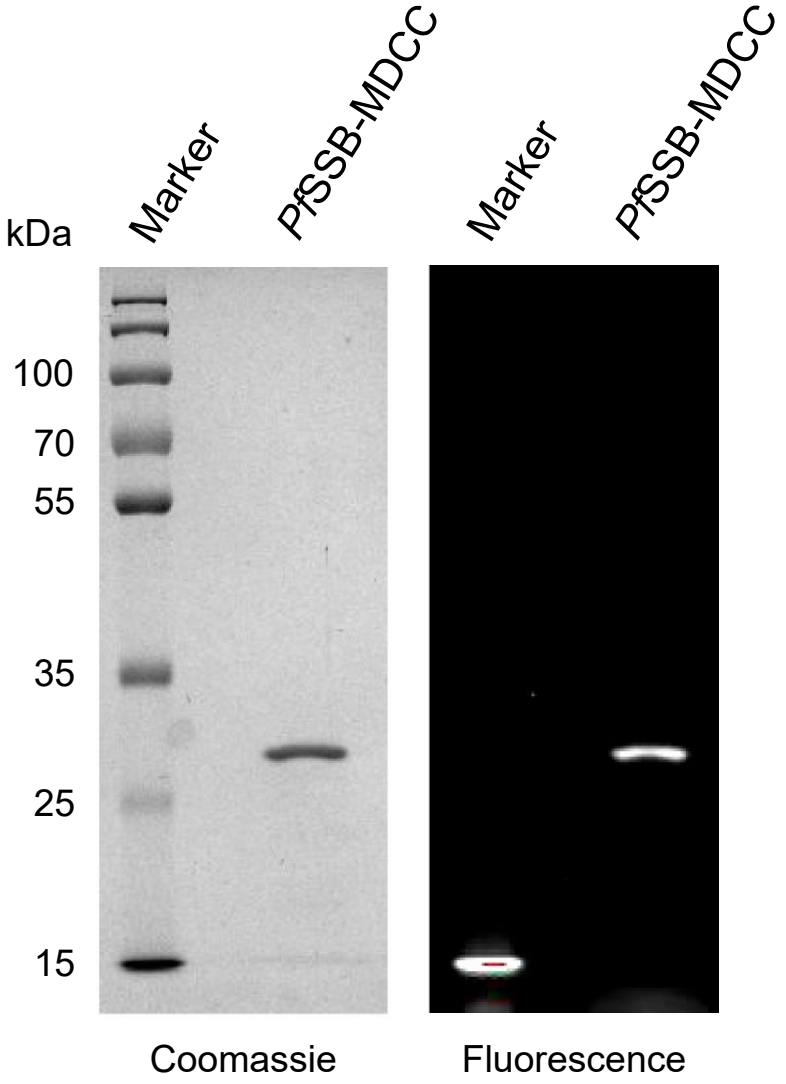



**Figure 21 | Labeling *Pf*SSB.** Purified and labeled *Pf*SSB-MDCC is imaged with coomassie staining and fluorescence. This data has been previously published.[76]

was complementary to the 5′ end of the other, 41 nucleotide strand. 1 μM concentrations of each strand were annealed in 20 mM Tris-HCl pH 8.0 and 50 mM NaCl.

*Stopped flow Instrumentation*

Stopped flow technique is discussed in Chapter 2.

## 3.4 Results and Discussion

### *3.4.1 PfSSB-MDCC Assays*

The usefulness of the *Pf*SSB probe is determined by how well it functions as a monitor, indicating when free ssDNA is available in a system. In order to investigate the quality of the *Pf*SSB probe, the probe must initially be characterized. Afterwards, the probe may be used to study different proteins in HR.

The labeled probe exhibits an excitation peak near 430 nm and an emittance peak near 475 nm. Therefore an excitation wavelength of 430 nm was used in fluorescence-measuring experiments (Fig. 22a). The probe must exhibit a change in fluorescence upon binding to ssDNA, otherwise it is not revealing when ssDNA is unbound. The probe by itself has a significant fluorescence signal in the absence of ssDNA (Fig. 22b). However, when mixed with $dT_{70}$ ssDNA, nearly a fourfold increase in fluorescence is seen at 475 nm. This change in signal intensity demonstrates that the probe is sensitive to available ssDNA.

Next an experiment was ran with the use of fluorimeter instrumentation with ratios of *Pf*SSB probe to $dT_{70}$ of 0 to 1.6 (Fig. 22c). As would be expected for a tetramer that binds 65 nucleotides of single stranded DNA, the fluorescence

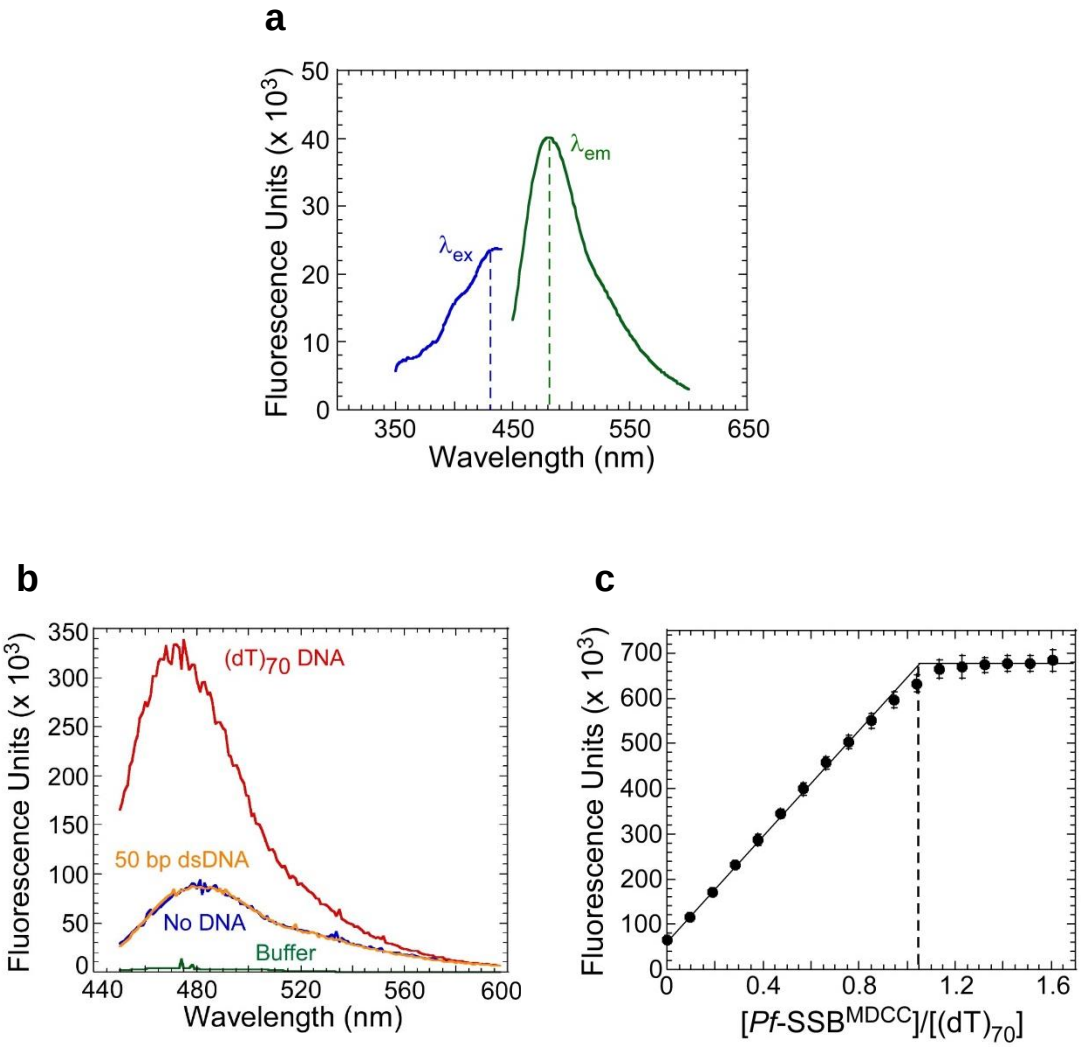



**Figure 22| *Pf*SSB probe fluorescence and binding ratio. a,** Observed are the excitation and emission wavelengths for the *Pf*SSB-MDCC probe. **b,** The difference in fluorescence intensity of the probe alone in buffer, and when it binds to a strand of ssdT$_{70}$ DNA, is nearly fourfold. **c,** The binding of *Pf*SSB-MDCC to ssdT$_{70}$ essentially ceases when there are equal parts of both probe and filament, displaying how one *Pf*SSB-MDCC tetramer binds to one sequence of dT$_{70}$. The data and images from 22a and 22c have been previously published in PLoS ONE.[76]

plateaued once a ratio of one unit of tetramer to one unit of $dT_{70}$ was reached. This displays that the probe binds only one length of $dT_{70}$ at a time, and that there is only one tetramer per unit $dT_{70}$.

As noted earlier, previous SSB probes show differing binding modes of the SSB tetramer, dependent upon NaCl concentration. Since *Pf*SSB has only one binding mode[75], the *Pf*SSB probe should bind to the same number of nucleotides of ssDNA at various NaCl concentrations. Figure 23a displays fluorimeter data of the probe, ran under the same conditions as done previously (Fig. 22c), but with NaCl concentrations of 0.01 M, 0.1 M, 0.50 M, and 1.0 M. These differing conditions all indicate the same binding mode, with the fluorescence signal saturating at a ratio of 1 *Pf*SSB probe to 1 length of $dT_{70}$ ssDNA. In Figure 23b, the binding stoichiometry from the experiment is more clearly displayed. In every condition of NaCl concentration, the binding stoichiometry remains near one.

In order to further characterize the probe, Stopped flow assays were performed with 20 nM *Pf*SSB in one sample, and increasing concentrations of $dT_{70}$ (25 nM, 50 nM, 100nM, 250 nM, and 400 nM) in the other (Fig. 24a.). To acquire the observed rate of *Pf*SSB binding to $dT_{70}$, the data was fit using a single exponential equation. The single exponential fit equation is as follows:

$$A_t = A_0 e^{-kt}$$

In this equation, $A_0$ is fluorescence (or bound $dT_{70}$) at time zero, $A_t$ is fluorescence (or bound $dT_{70}$) at time t, k is the observed rate constant and t is time. Graphing the observed rate constant against the concentration of $dT_{70}$

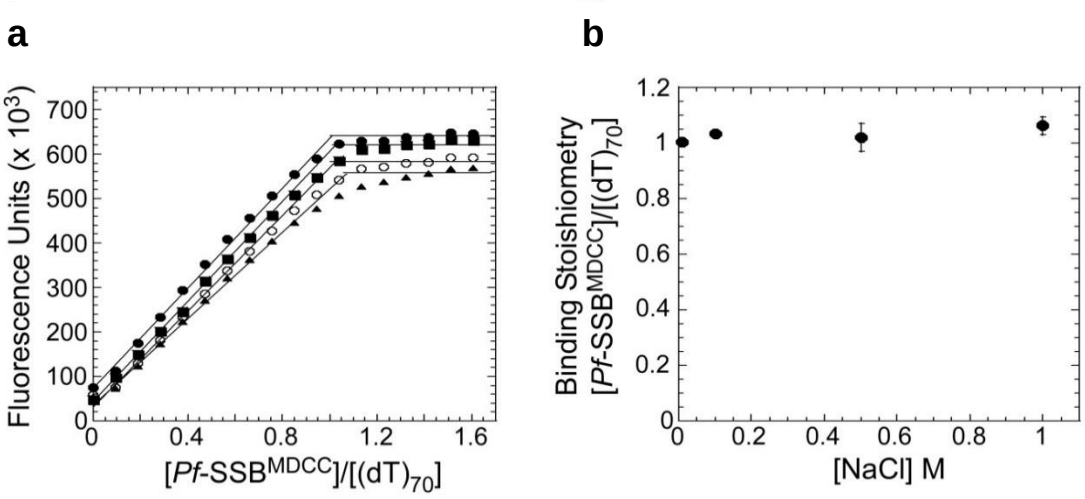



**Figure 23 | *Pf*SSB-MDCC exhibits a single binding mode. a,** For differencing concentrations of NaCl, fluorescence is observed against the number of *Pf*SSB tetramer probes, divided by the number of $dT_{70}$ molecules. **b,** For all four NaCl concentrations, a plateau in fluorescence is observed once the number of *Pf*SSB tetramer probes, divided by the number of $dT_{70}$, equals roughly one. This indicates that PfSSB exhibits only one binding mode. The data and images found here have been previously published.[76]

results in a linear plot (Fig. 24b). The slope of this plot provides the rate of association, 2.6 x 10^8 $M^{-1}s^{-1}$. This demonstrates that the probe binds ssDNA quickly, fulfilling a necessary requirement of sensitivity for free ssDNA. To be functional in studying HR, the probe must bind ssDNA at a fast rate, out competing potential ssDNA-reactive molecules, including the scenario where the ssDNA self-anneals. As SSB works to protect free ssDNA from damage and self-annealing in nature, the tetramer can bind quickly enough to prevent these potentially hindersome outcomes from happening.

Now that the probe has been characterized, it may be utilized as a tool in the study of HR proteins. One such experiment was performed, with the use of Stopped flow instrumentation, to monitor the disassembly of the Rad51-ssDNA

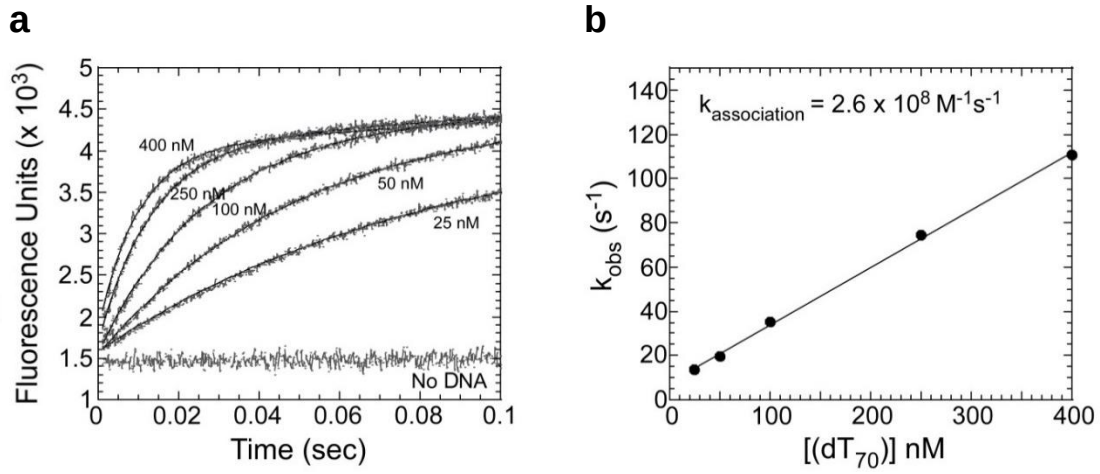



**Figure 24| Determining the rate of association of *Pf*SSB-MDCC. a,** The initial rate of *Pf*SSB binding to $dT_{70}$ DNA is observed at differing DNA concentrations. **b,** Fitting each concentration of $dT_{70}$ to a single exponential provides the observed rate. Plotting the observed rates against substrate concentration gives a linear plot, the slope of which provides a rate of association, 2.6 x $10^8$ $M^{-1}s^{-1}$. These images have been previously published in PLoS ONE.[76]

filament by the SRS2 anti-recombinase protein. The initial condition was performed by mixing 75 nM of *Pf*SSB probe with 3uM M13mp18 plasmid ssDNA (New England Biolabs). Rad51 Clearing Buffer (25 mM Tris-HCl pH 7.7, 1 mM DTT, 5% glycerol, 10 mM $MgCl_2$, and 50 mM NaCl) was used in all trials. As expected, a quick increase in fluorescence signal was seen as the probe bound residues of free ssDNA (Fig. 25a).

In the next condition, the probe was mixed with M13mp18 plasmid, which had been pre-incubated with 1.5 μM Rad51 and 2.85 μM ATP (Fig. 25b). As the Rad51 filament forms and is already bound to the ssDNA, no significant increase in signal was observed upon mixing. As the probe is not signaling for ssDNA binding, the Rad51 protein is not being replaced by the *Pf*SSB. The observation

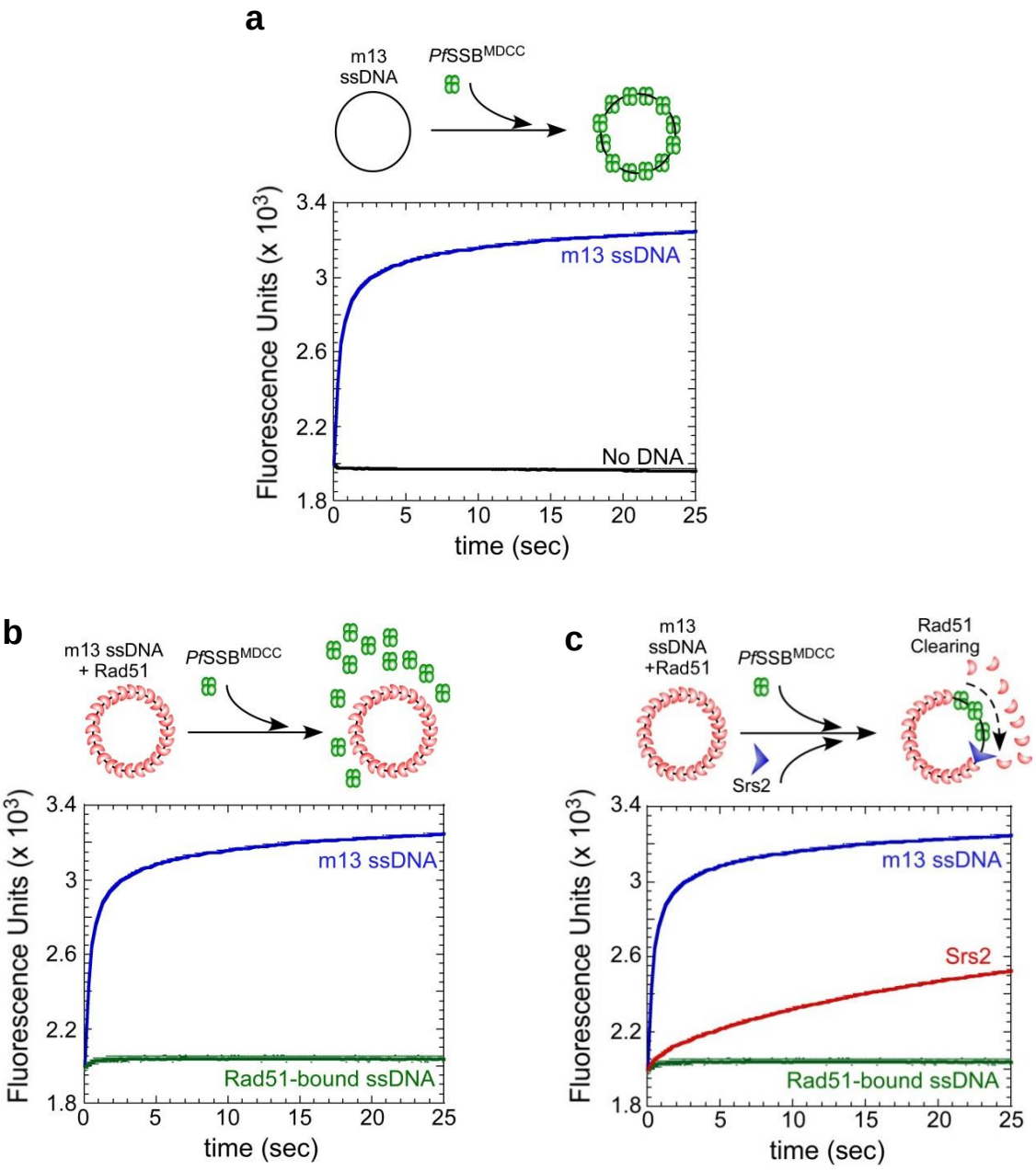



**Figure 25 | SRS2 Rad51 clearing experiment. a,** The probe binds quickly to m13ss plasmid DNA. **b,** A pre-formed Rad51 filament on the M13mp18 plasmid ssDNA is not replaced by *Pf*SSB. **c,** When SRS2 anti-recombinase protein is added along with the probe to the preformed Rad51-DNA filament, Rad51 is removed by SRS2 and replaced with the *Pf*SSB probe on the plasmid. These images have been previously published.[76]

that *Pf*SSB does not remove the Rad51 filament is essential, otherwise the probe could not act as a reliable reporter on how other HR proteins disassemble Rad51 from ssDNA.

Finally, (now that it is known that the probe does not, at least on its own, remove the Rad51 filament), a condition was ran to study the activity of the SRS2 anti-recombinase. In this trial, the *Pf*SSB probe, as well as 200 nM SRS2, were mixed with the pre-incubated Rad51, ATP, and M13mp18 ssDNA plasmid (Fig. 25c). A steady increase in fluorescence intensity is seen as SRS2 moves along the ssDNA plasmid, removing the Rad51 filament. The probe, with the previously determined rate of association of $2.6 \times 10^8$ $M^{-1}s^{-1}$, is able to bind to the ssDNA after the dissociation of Rad51. SRS2 has been shown to remove Rad51 from ssDNA at a rate of approximately 12 nucleotides $s^{-1}$ [37].

The more time SRS2 is allowed to remove Rad51, the more signal intensity, and thus the more binding of *Pf*SSB, is observed. This experiment provides supporting evidence for SRS2's role as an anti-recombinase in HR, and that the *Pf*SSB probe can be used to study the disassociation of Rad51 bound to ssDNA. Hence the probe is seen as a successful tool for the study of HR proteins.

### *3.4.2 SRS2 DNA Unwinding*

The next experiment performed with *Pf*SSB focused on the helicase activity of SRS2, rather than its function as an anti-recombinase. A model of the experiment can be seen in Figure 26. As previously noted, helicase activity is a

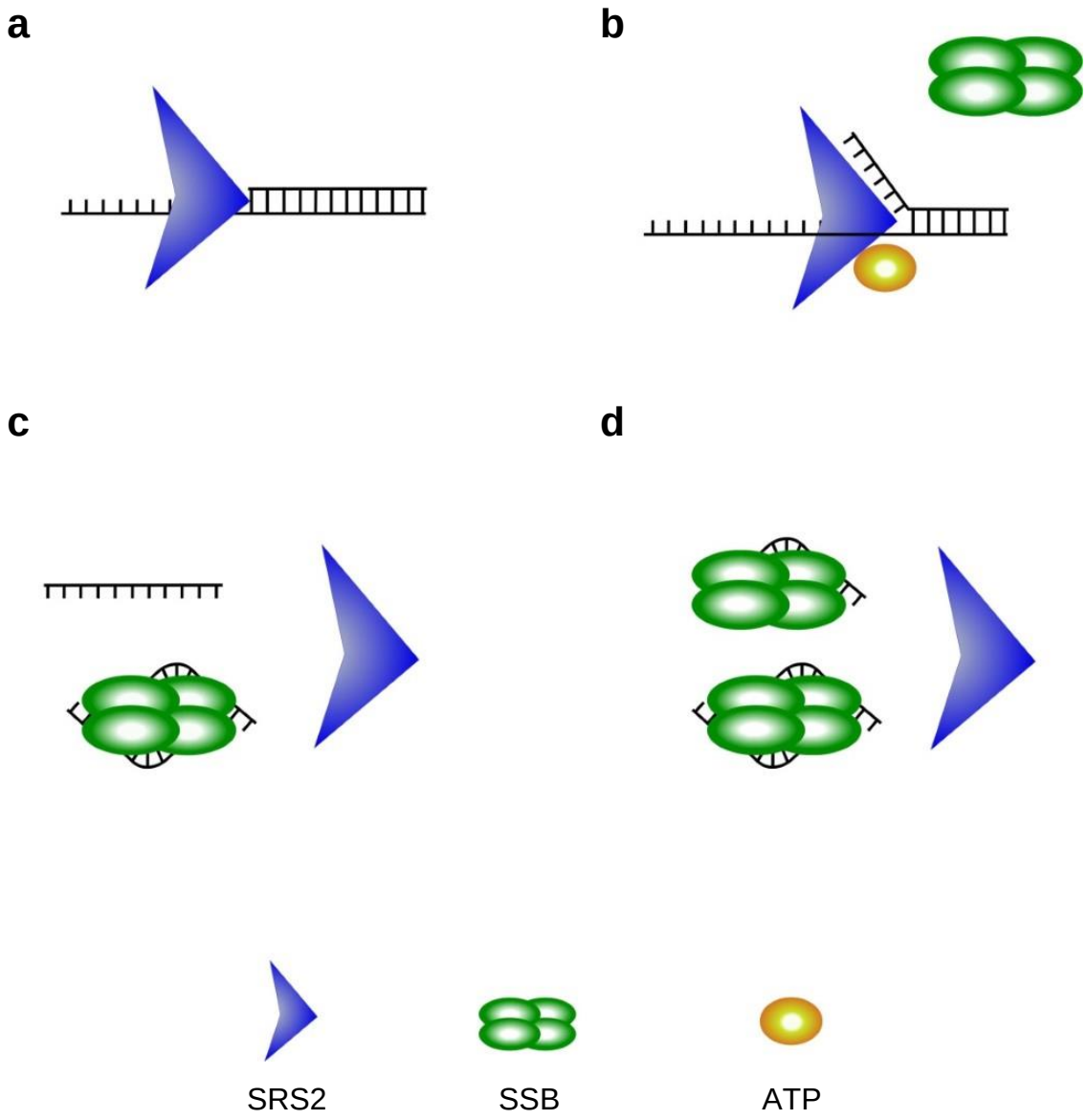



**Figure 26 | A model for the SRS2 DNA unwinding experiment. a,** SRS2 is pre-incubated with DNA, with 25 double stranded nucleotides attached to a 16 nucleotide single-stranded overhang. The SRS2 loads onto the ssDNA. **b,** Upon mixing with excess ATP and two-fold excess PfSSB, SRS2 begins to unwind the dsDNA. **c,** The PfSSB probe binds to the newly freed ssDNA, and undergoes a change in fluorescence intensity. **d,** Due to the two-fold excess of PfSSB, both of the resulting ssDNA molecules can be bound.

vital component of homologous recombination. As PfSSB binds to ssDNA, it can be used to monitor when a helicase unzips dsDNA, resulting with two ssDNA strands that are free to bind to *Pf*SSB. With the use of Stopped flow instrumentation, a solution of 160 nM wild type SRS2 protein, 2.85 μM ATP and 40 nM annealed dsDNA, with a 3′ ssDNA overhang, was mixed with a solution of 80 nM *Pf*SSB. DNA Unwinding Buffer (50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, and 50 mM NaCl) was used in all conditions. SRS2 binds to ssDNA overhangs[77], and translocates with 3′ to 5′ directionality[16].

These conditions were repeated once with a no ATP control, and once more with an SRS2 containing a DK-AA mutation in its 2B domain. This mutated protein was used to test if the changed residues will alter SRS2 DNA unwinding activity, as was observed in a homologous helicase, UVRD[78].

In the no ATP control, only a very slight increase in fluorescence intensity was observed (Fig. 27). This indicates that SRS2 does not unwind the dsDNA in the absence of ATP, and that the *Pf*SSB probe does not work to displace the SRS2 from its position on the single-stranded overhang of the DNA. Some increase in fluorescence can be accounted for as SRS2 slowly dissociates from the single stranded overhang, which is then bound by *Pf*SSB.

For both trials including ATP, the SRS2, utilizing ATP as an energy source, unwinds the dsDNA. This results in the large, quick increase in fluorescence by the *Pf*SSB probe (Fig. 27). As both versions of the protein show similar cleavage data, the results of this experiment indicate that the mutation

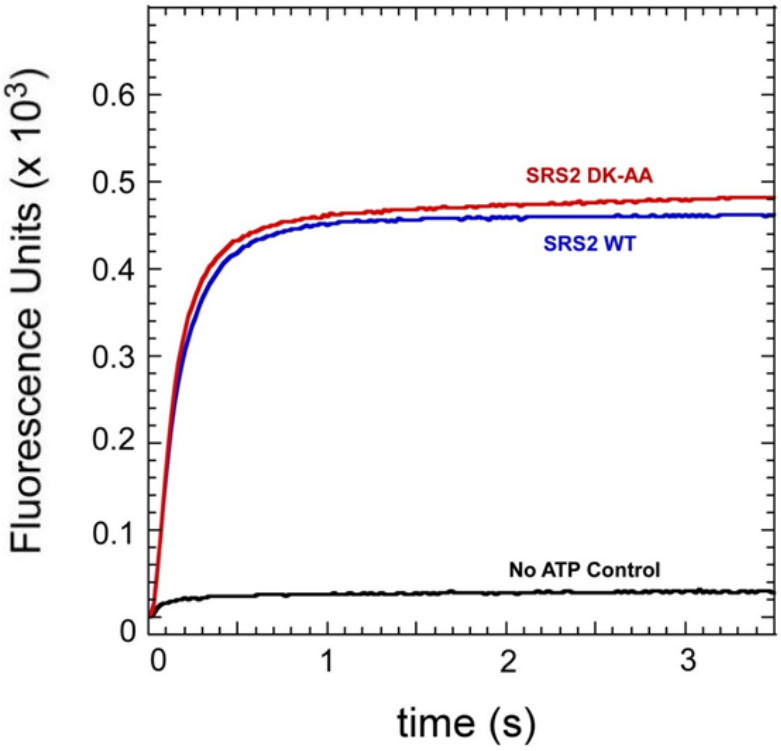



**Figure 27 | SRS2 DNA unwinding data.** Displayed are the results from a Stopped flow experiment. One of the samples to be mixed contains SRS2, pre-loaded onto dsDNA with a ssDNA overhang, while the other the contains the *Pf*SSB probe and ATP. Two variants of the SRS2 protein were investigated, the wild type and the AA-DK 2B domain mutant.

does not alter the DNA unwinding activity of SRS2. While no significant difference was found in helicase activity between the two SRS2 proteins, this experiment does demonstrate that the labeled tetramer works as a probe for helicase activity, which is a crucial component of homologous recombination.

### 3.5 Future Work

This probe could function in further studies on HR mediator proteins, especially those which either promote or hinder the removal of the Rad51

filament, to indicate how fast ssDNA is becoming available in a system. Additionally, helicases, and other proteins that effect helicase function, can be investigated utilizing the fluorescently labeled *Pf*SSB. As the probe does not change binding modes, experiments can be ran at various concentrations of NaCl.

The assays performed here characterize the probe, and provide examples of how it can both function to monitor the removal of Rad51 from ssDNA, as well as the unwinding of dsDNA. Working to monitor Rad51, as well as the HR protein SRS2, demonstrates that this probe can function as a tool in the further study of HR.

# CHAPTER IV

# SUMMARY

Rad51, the primary machinery for HR, is critical for genome maintenance. The up-regulation of Rad51 and HR can lead to excessive genetic rearrangements, causing genomic instability. Likewise, down-regulation can lead to an accumulation of double-strand breaks. Both of these scenarios can be harmful to the cell. The proper balance of HR and Rad51 is key, various cancers have demonstrated either up or down regulation of Rad51. This stresses the importance of studying the proteins involved in mediating the Rad51-DNA filament in HR, particularly where the specific function and mechanism of many Rad51 filament mediator proteins is largely unknown.

To increase the understanding of these mediator proteins, this work has developed a toolkit of protein-probes. These proteins are directly attached to fluorophores at specific amino acid residues. While a common technique in studying HR proteins involves the use of DNA labeled with a terminal fluorophore, labeled proteins can provide a signal at essentially every occurrence of a DNA-protein interaction. This provides greater experimental detail in the study of how different proteins effect the Rad51-DNA filament.

In the making of this tool kit, probes were first developed from *Sc*Rad51. This was completed in accordance with **Aim 1*:*** Create a probe for the Rad51-DNA filament by incorporating the unnatural amino acid 4-amino-l-phenylalanine in place of specific amino acid residues in *Saccharomyces cerevisiae* Rad51.

Successful incorporation of 4AZP was performed at various positions of Rad51's loops L1 and L2. The Rad51-4AZP-fluropohre Stopped flow assay preliminary results indicate a change in fluorescence emission intensity upon protein-ssDNA filament formation with probes located at positions 290, 291, and 337. This is in line with **Aim 1a:** Specifically investigate positions on loops L1 and L2 of *Sc*Rad51 which, following labeling with an alkyne fluorophore, will result in fluorescently labeled protein which displays a change in fluorescence emission intensity upon filament formation.

Further locations on loops L1 and L2 can be investigated in future work to find a greater signal change upon DNA binding. Future work can also optimize these probes through increased labeling efficiently. Different click fluorophores will be tested at various temperatures and durations.

The size and shape of individual fluorophores could also impact the change in fluorescence emission intensity upon filament formation, as well as effect Rad51 activity levels. While the labeled probes may undergo an observable change in fluorescence emission intensity upon binding to DNA, protein activity levels should not be altered for the probes to accurately report Rad51 filament dynamics. If a fluorophore is too bulky and inhibits filament formation, the probe would not perform activity at the same level as the wild type.

While it has been displayed that 4AZP incorporation does not alter filament formation on Cy3 labeled $dT_{79}$ ssDNA, further work will be performed to measure the activity levels of the Rad51 probe with a fluorophore bound. This will

be measured with the utilization of electrophoretic mobility shift assays (EMSA). Here, Rad51 probes and wild type Rad51 will be evaluated with the use of DNA oligomers. The first oligomer will be a short, single strand of DNA radiolabeled with 5′ phosphorous-32 (P-32). P-32 is a radioactive isotope, undergoing beta decay to convert to sulfur-32.

Another oligomer will consist of a large, circular plasmid, with a section of sequence complimentary to the first small strand. This plasmid will be annealed to another short strand with the exact same sequence as the radiolabeled DNA, and without the P-32 label. Rad51 will perform strand exchange activity with the two short oligomers, replacing the oligomer attached to the plasmid with the P-32 version. The reaction will be quenched at given time points, and samples ran on a polyacrylamide gel. If strand exchange has occurred, the P-32 label can be attached to the large plasmid, appearing in an upper (greater mass) band on the gel. If strand exchange activity is lowered for the Rad51-4AZP probes, a greater amount of the P-32 label will be seen in a lower (lesser mass) band for the same time points. This corresponds to the short P-32 sequence of DNA that has not been exchanged onto the plasmid. By performing the assay with differing concentrations of substrate DNA, rates for strand exchange activity could be obtained for the wild type and 4AZP incorporated proteins.

After labeling efficiency has been optimized and activity levels for the probes have been determined to be equivalent to the wild type, further experiments may be performed. Ideally, the probes with fluorophores positioned

on loops L1 and L2 can be used to monitor the effects of HR mediator proteins on the Rad51 filament. For these experiments to proceed, the probes will be mixed with DNA, along with a given mediator protein such as Rad52.

Rad52 promotes the replacement of RPA with Rad51 on DNA[36]. A future experiment could be performed by mixing DNA already bound by RPA with Rad51 and Rad52 in a Stopped flow apparatus. A change in fluorescence emission intensity should be observed as the Rad51 replaces RPA and forms a filament with DNA. This could potentially reveal a rate for the replacement of RPA by Rad51 in the presence of Rad52.

Additional experiments could be performed in a similar manner, but instead with Rad51 pre-incubated on DNA and mixed with RPA and SRS2. As SRS2 disassociates the Rad51 filament, allowing RPA to bind in its place, a rate for Rad51 filament removal by SRS2 could be obtained by performing the experiment at various Rad51-DNA filament concentrations.

While current techniques utilize short strands of DNA with a terminal fluorophore, large lengths of DNA, such as plasmids, can be studied with these probes. This is significant, as Rad51 filaments form on long strands of DNA in nature when they repair the double-strand breaks of a chromosome. Instead of having an end reporter of the Rad51-DNA interaction, reporters are present at every location of interaction between DNA and Rad51. With labeled DNA, any nearby protein may cause a change in fluorescence intensity. By labeling

proteins instead of nucleotides, information can be gathered on which specific protein is bound.

A probe has been positioned near the N-terminus of Rad51 (16-4AZP), in the pursuance of **Aim 1b:** Develop a UAA probe with a fluorescent label near the N-terminus of *Sc*Rad51, to be used for future FRET and single molecule designs. The same techniques will be performed to optimize labeling efficiency and check the level of activity. Next, FRET experiments could be performed where 50% of the protein is labeled with one member of a FRET pair, while 50% of the protein is labeled with the other. This experiment would display when a filament has been formed, with a high FRET signal, as the FRET pair fluorophores are close enough to each other in the filament to undergo FRET. Likewise, filament disassembly can be observed by a decrease in FRET signal. This experimental design is further expounded upon in Chapter 2.

In addition to the FRET experimentation, future work can also utilize this probe for single molecule studies. As the position of the fluorophore is on the n-terminal domain, away from the sites of DNA and ATP binding, this probe should not interfere with protein activity, (though this will still be checked with the EMSA discussed previously). If the labeled probe is allowed to form on an isolated strand of DNA, the use of a total internal reflection fluorescence (TIRF) microscope might provide high detail images of the Rad51-DNA filament. This could provide insights into nucleation and the exchange of RPA for Rad51, as the

images would spatially display were the labeled probes are situated during the different steps of HR.

Additional success has been seen with the development of a fluorescently labeled probe from *Plasmodium falciparum* SSB. As this probe wraps DNA very tightly, it can act as a reporter for when free single stranded DNA becomes available with a near fourfold increase in fluorescence intensity.

The *Pf*SSB probe is simple to make; no amino acid residues must be changed via mutation in order to produce the probe. This leaves the protein unaltered, save for the addition of a fluorophore at a lone cysteine residue. *Pf*SSB, unlike *Ec*SSB, displays only one salt-dependent binding mode where it binds roughly one unit of $dT_{70}$ per unit tetramer. This provides an advantage in experiments, as salt conditions do not have to be regimented to insure the quality of results.

The binding constant of this probe was calculated to be $2.6 \times 10^8$ $M^{-1}s^{-1}$, demonstrating that ssDNA will be bound very quickly in reactions. Thus, this probe can be useful in Stopped flow experimentation, where millisecond-scale measurements are taken upon mixing two solutions. Such is the case in our experiment where a pre-formed Rad51 filament is cleared by SRS2 from ssDNA, allowing the SSB probe to bind in its place and exhibit a corresponding increase in fluorescence signal. We have also displayed that the probe can be used to monitor helicase activity. This probe works in accordance with **Aim 2:** With Plasmodium falciparum single strand DNA binding protein (SSB), develop a

probe which, when labeled with a maleimide fluorophore, will act as a signal for unbound, single stranded DNA (ssDNA) by undergoing a change in fluorescence emission intensity upon binding.

While this work utilized the *Pf*SSB probe for studying the anti-recombinase protein SRS2, other HR proteins could also be investigated. Future experimentation with the *Pf*SSB probe could study the Shu complex, which promotes homologous recombination. The ability of the Shu complex to potentially alter the rate of Rad51 filament removal by anti-recombinase proteins such as SRS2 could be examined.

In this work, protein-fluorophore probes have been developed for purpose of studying Rad51 filament dynamics. Advancing the current techniques available to investigate homologous recombination can help further the current understanding of the interactions between Rad51 and HR mediator proteins. The association of cancers with altered HR levels makes the topic all the more relevant. Revealing the mechanisms and function of these mediators could help indicate directions for future therapeutic approaches in the treatment of cancers and associated genetic disorders.

# REFERENCES


1. Watson, J. D. & Crick, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **171,** 737–738 (1953).
2. Schärer, O. D. Chemistry and Biology of DNA Repair. *Angew. Chem. Int. Ed.* **42,** 2946–2974 (2003).
3. Filippo, J. S., Sung, P. & Klein, H. Mechanism of Eukaryotic Homologous Recombination. *Annu. Rev. Biochem.* **77,** 229–257 (2008).
4. Li, X. & Heyer, W.-D. Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res.* **18,** 99–113 (2008).
5. Krogh, B. O. & Symington, L. S. Recombination Proteins in Yeast. *Annu. Rev. Genet.* **38,** 233–271 (2004).
6. Garinis, G. A., van der Horst, G. T. J., Vijg, J. & Hoeijmakers, J. H. J. DNA damage and ageing: new-age ideas for an age-old problem. *Nat. Cell Biol.* **10,** 1241–1247 (2008).
7. Kowalczykowski, S. C., Dixon, D. A., Eggleston, A. K., Lauder, S. D. & Rehrauer, W. M. Biochemistry of homologous recombination in Escherichia coli. *Microbiol. Rev.* **58,** 401–465 (1994).
8. Namsaraev, E. & Berg, P. Characterization of strand exchange activity of yeast Rad51 protein. *Mol. Cell. Biol.* **17,** 5359–5368 (1997).
9. Sung, P. Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast RAD51 protein. *Science* **265,** 1241–1243 (1994).

10. Sung, P. & Klein, H. Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat. Rev. Mol. Cell Biol.* **7,** 739–750 (2006).

11. Chapman, J. R., Taylor, M. R. G. & Boulton, S. J. Playing the End Game: DNA Double-Strand Break Repair Pathway Choice. *Mol. Cell* **47,** 497–510 (2012).

12. Moynahan, M. E. & Jasin, M. Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. *Nat. Rev. Mol. Cell Biol.* **11,** 196–207 (2010).

13. Cromie, G. A. & Leach, D. R. F. Control of Crossing Over. *Mol. Cell* **6,** 815–826 (2000).

14. Krejci, L. *et al.* Interaction with Rad51 Is Indispensable for Recombination Mediator Function of Rad52. *J. Biol. Chem.* **277,** 40132–40141 (2002).

15. Krejci, L., Altmannova, V., Spirek, M. & Zhao, X. Homologous recombination and its regulation. *Nucleic Acids Res.* **40,** 5795–5818 (2012).

16. Liu, J. *et al.* Rad51 paralogs Rad55-Rad57 balance the anti-recombinase Srs2 in Rad51 filament formation. *Nature* **479,** 245–248 (2011).

17. Gaines, W. A. *et al.* Promotion of presynaptic filament assembly by the ensemble of S. cerevisiae Rad51 paralogues with Rad52. *Nat. Commun.* **6,** 7834 (2015).

18. Li, X. & Heyer, W.-D. RAD54 controls access to the invading 3′-OH end after RAD51-mediated DNA strand invasion in homologous recombination in Saccharomyces cerevisiae. *Nucleic Acids Res.* **37,** 638–646 (2009).

19. Sengupta, S. & Harris, C. C. p53: traffic cop at the crossroads of DNA repair and recombination. *Nat. Rev. Mol. Cell Biol.* **6,** 44–55 (2005).

20. Linke, S. P. *et al.* p53 interacts with hRAD51 and hRAD54, and directly modulates homologous recombination. *Cancer Res.* **63,** 2596–2605 (2003).

21. Marmorstein, L. Y., Ouchi, T. & Aaronson, S. A. The BRCA2 gene product functionally interacts with p53 and RAD51. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 13869–13874 (1998).

22. Krejci, L. *et al.* DNA helicase Srs2 disrupts the Rad51 presynaptic filament. *Nature* **423,** 305–309 (2003).

23. Prakash, R. *et al.* Yeast Mph1 helicase dissociates Rad51-made D-loops: implications for crossover control in mitotic recombination. *Genes Dev.* **23,** 67–79 (2009).

24. Mitchel, K., Lehner, K. & Jinks-Robertson, S. Heteroduplex DNA Position Defines the Roles of the Sgs1, Srs2, and Mph1 Helicases in Promoting Distinct Recombination Outcomes. *PLOS Genet* **9,** e1003340 (2013).

25. Ford, D., Easton, D. F., Bishop, D. T., Narod, S. A. & Goldgar, D. E. Risks of cancer in BRCA1-mutation carriers. Breast Cancer Linkage Consortium. *Lancet Lond. Engl.* **343,** 692–695 (1994).

26. Poh, W. *et al.* Epigenetic Silencing of BRCA1 Is Linked to Homologous Recombination Repair Defects and Elevated Mir-155 Expression in Myeloid Neoplasms. *Blood* **124,** 3525–3525 (2014).

27. He, X. & Zhang, P. Serine/arginine-rich splicing factor 3 (SRSF3) regulates homologous recombination-mediated DNA repair. *Mol. Cancer* **14,** (2015).

28. Zhao, W. *et al.* Promotion of BRCA2-Dependent Homologous Recombination by DSS1 via RPA Targeting and DNA Mimicry. *Mol. Cell* **59,** 176–187 (2015).

29. Kidane, D., Ayora, S., Sweasy, J. B., Graumann, P. L. & Alonso, J. C. The cell pole: the site of cross talk between the DNA uptake and genetic recombination machinery. *Crit. Rev. Biochem. Mol. Biol.* **47,** 531–555 (2012).

30. Hine, C. M., Seluanov, A. & Gorbunova, V. Use of the Rad51 promoter for targeted anti-cancer therapy. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 20810–20815 (2008).

31. Hine, C. M., Seluanov, A. & Gorbunova, V. Rad51 promoter-targeted gene therapy is effective for in vivo visualization and treatment of cancer. *Mol. Ther. J. Am. Soc. Gene Ther.* **20,** 347–355 (2012).

32. Yoshikawa, K. *et al.* Abnormal expression of BRCA1 and BRCA1-interactive DNA-repair proteins in breast carcinomas. *Int. J. Cancer* **88,** 28–36 (2000).

33. Bennett, B. T. & Knight, K. L. Cellular localization of human Rad51C and regulation of ubiquitin-mediated proteolysis of Rad51. *J. Cell. Biochem.* **96,** 1095–1109 (2005).

34. Binz, S. K., Sheehan, A. M. & Wold, M. S. Replication Protein A phosphorylation and the cellular response to DNA damage. *DNA Repair* **3,** 1015–1024 (2004).

35. Maréchal, A. & Zou, L. RPA-coated single-stranded DNA as a platform for post-translational modifications in the DNA damage response. *Cell Res.* **25,** 9–23 (2015).

36. Sugiyama, T. & Kowalczykowski, S. C. Rad52 Protein Associates with Replication Protein A (RPA)-Single-stranded DNA to Accelerate Rad51-mediated Displacement of RPA and Presynaptic Complex Formation. *J. Biol. Chem.* **277,** 31663–31672 (2002).

37. Antony, E. *et al.* Srs2 Disassembles Rad51 Filaments by a Protein-Protein Interaction Triggering ATP Turnover and Dissociation of Rad51 from DNA. *Mol. Cell* **35,** 105–115 (2009).

38. Lytle, A. K. *et al.* Context-Dependent Remodeling of Rad51–DNA Complexes by Srs2 Is Mediated by a Specific Protein–Protein Interaction. *J. Mol. Biol.* **426,** 1883–1897 (2014).

39. Bermek, E. *Mechanisms of Protein Synthesis: Structure-Function Relations, Control Mechanisms, and Evolutionary Aspects*. (Springer Science & Business Media, 2012).

40. Krejci, L., Damborsky, J., Thomsen, B., Duno, M. & Bendixen, C. Molecular Dissection of Interactions between Rad51 and Members of the Recombination-Repair Group. *Mol. Cell. Biol.* **21,** 966–976 (2001).

41. Budke, B. *et al.* RI-1: a chemical inhibitor of RAD51 that disrupts homologous recombination in human cells. *Nucleic Acids Res.* **40,** 7347–7357 (2012).

42. Wang, J., Xie, J. & Schultz, P. G. A Genetically Encoded Fluorescent Amino Acid. *J. Am. Chem. Soc.* **128,** 8738–8739 (2006).

43. Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C. & Schultz, P. G. A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* **244,** 182–188 (1989).

44. Hoesl, M. G. & Budisa, N. Expanding and engineering the genetic code in a single expression experiment. *Chembiochem Eur. J. Chem. Biol.* **12,** 552–555 (2011).

45. Mukai, T. *et al.* Codon reassignment in the Escherichia coli genetic code. *Nucleic Acids Res.* **38,** 8188–8195 (2010).

46. Huang, Y. *et al.* A convenient method for genetic incorporation of multiple noncanonical amino acids into one protein in Escherichia coli. *Mol. Biosyst.* **6,** 683–686 (2010).

47. Schneider, T. *et al.* Dissecting ubiquitin signaling with linkage-defined and protease resistant ubiquitin chains. *Angew. Chem. Int. Ed Engl.* **53,** 12925–12929 (2014).

48. Chen, Z., Yang, H. & Pavletich, N. P. Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature* **453,** 489–484 (2008).

49. Chen, J., Villanueva, N., Rould, M. A. & Morrical, S. W. Insights into the mechanism of Rad51 recombinase from the structure and properties of a filament interface mutant. *Nucleic Acids Res.* **38,** 4889–4906 (2010).

50. Conway, A. B. *et al.* Crystal structure of a Rad51 filament. *Nat. Struct. Mol. Biol.* **11,** 791–796 (2004).

51. Yang, J., Zhang, R., Radford, D. C. & Kopeček, J. FRET-trackable biodegradable HPMA copolymer-epirubicin conjugates for ovarian carcinoma therapy. *J. Controlled Release* **218,** 36–44 (2015).

52. Kim, Y., Kim, S. H., Ferracane, D., Katzenellenbogen, J. A. & Schroeder, C. M. Specific Labeling of Zinc Finger Proteins using Noncanonical Amino Acids and Copper-Free Click Chemistry. *Bioconjug. Chem.* **23,** 1891–1901 (2012).

53. Yuan, F. *et al.* Use of a novel Förster resonance energy transfer method to identify locations of site-bound metal ions in the U2–U6 snRNA complex. *Nucleic Acids Res.* **35,** 2833–2845 (2007).

54. Hammill, J. T., Miyake-Stoner, S., Hazen, J. L., Jackson, J. C. & Mehl, R. A. Preparation of site-specifically labeled fluorinated proteins for 19F-NMR structural characterization. *Nat. Protoc.* **2,** 2601–2607 (2007).

55. Flickinger, M. C. *Upstream Industrial Biotechnology, 2 Volume Set*. (John Wiley & Sons, 2013).

56. Lewinska, A., Wnuk, M., Slota, E. & Bartosz, G. Total Anti-Oxidant Capacity of Cell Culture Media. *Clin. Exp. Pharmacol. Physiol.* **34,** 781–786 (2007).

57. Berg, J. M. *et al. Biochemistry*. (W H Freeman, 2002).

58. Desai, T. A. & Rao, C. V. Regulation of Arabinose and Xylose Metabolism in Escherichia coli. *Appl. Environ. Microbiol.* **76,** 1524–1532 (2010).

59. Magasanik, B. Catabolite Repression. *Cold Spring Harb. Symp. Quant. Biol.* **26,** 249–256 (1961).

60. Görke, B. & Stülke, J. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat. Rev. Microbiol.* **6,** 613–624 (2008).

61. Duong-Ly, K. C. & Gabelli, S. B. in *Methods in Enzymology* (ed. Lorsch, J.) **541,** 85–94 (Academic Press, 2014).

62. Englard, S. & Seifter, S. in *Methods in Enzymology* (ed. Deutscher, M. P.) **182,** 285–300 (Academic Press, 1990).

63. Kolb, H. C., Finn, M. G. & Sharpless, K. B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chem. Int. Ed.* **40,** 2004–2021 (2001).

64. Ravikumar, Y., Nadarajan, S. P., Hyeon Yoo, T., Lee, C. & Yun, H. Unnatural amino acid mutagenesis-based enzyme engineering. *Trends Biotechnol.* **33,** 462–470 (2015).

65. Kelly, T. J., Simancek, P. & Brush, G. S. Identification and characterization of a single-stranded DNA-binding protein from the archaeon Methanococcus jannaschii. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 14634–14639 (1998).

66. Shereda, R. D., Kozlov, A. G., Lohman, T. M., Cox, M. M. & Keck, J. L. SSB as an Organizer/Mobilizer of Genome Maintenance Complexes. *Crit. Rev. Biochem. Mol. Biol.* **43,** 289–318 (2008).

67. Weiner, J. H., Bertsch, L. L. & Kornberg, A. The deoxyribonucleic acid unwinding protein of Escherichia coli. Properties and functions in replication. *J. Biol. Chem.* **250,** 1972–1980 (1975).

68. T M Lohman & Ferrari, and M. E. Escherichia Coli Single-Stranded DNA-Binding Protein: Multiple DNA-Binding Modes and Cooperativities. *Annu. Rev. Biochem.* **63,** 527–570 (1994).

69. Zhang, W. *et al.* Directional Loading and Stimulation of PcrA Helicase by the Replication Initiator Protein RepD. *J. Mol. Biol.* **371,** 336–348 (2007).

70. Dillingham, M. S. *et al.* Fluorescent Single-Stranded DNA Binding Protein as a Probe for Sensitive, Real-Time Assays of Helicase Activity. *Biophys. J.* **95,** 3330–3339 (2008).

71. Bobst, E. V., Bobst, A. M., Perrino, F. W., Meyer, R. R. & Rein, D. C. Variability in the nucleic acid binding site size and the amount of single-

stranded DNA-binding protein in Escherichia coli. *FEBS Lett.* **181,** 133–137 (1985).

72. Lohman, T. M. & Overman, L. B. Two binding modes in Escherichia coli single strand binding protein-single stranded DNA complexes. Modulation by NaCl concentration. *J. Biol. Chem.* **260,** 3594–3603 (1985).

73. Green, M. *et al.* Engineering a reagentless biosensor for single-stranded DNA to measure real-time helicase activity in Bacillus. *Biosens. Bioelectron.* **61,** 579–586 (2014).

74. Antony, E., Weiland, E. A., Korolev, S. & Lohman, T. M. Plasmodium falciparum SSB Tetramer Wraps Single-Stranded DNA with Similar Topology but Opposite Polarity to E. coli SSB. *J. Mol. Biol.* **420,** 269–283 (2012).

75. Antony, E., Kozlov, A. G., Nguyen, B. & Lohman, T. M. Plasmodium falciparum SSB Tetramer Binds Single-Stranded DNA Only in a Fully Wrapped Mode. *J. Mol. Biol.* **420,** 284–295 (2012).

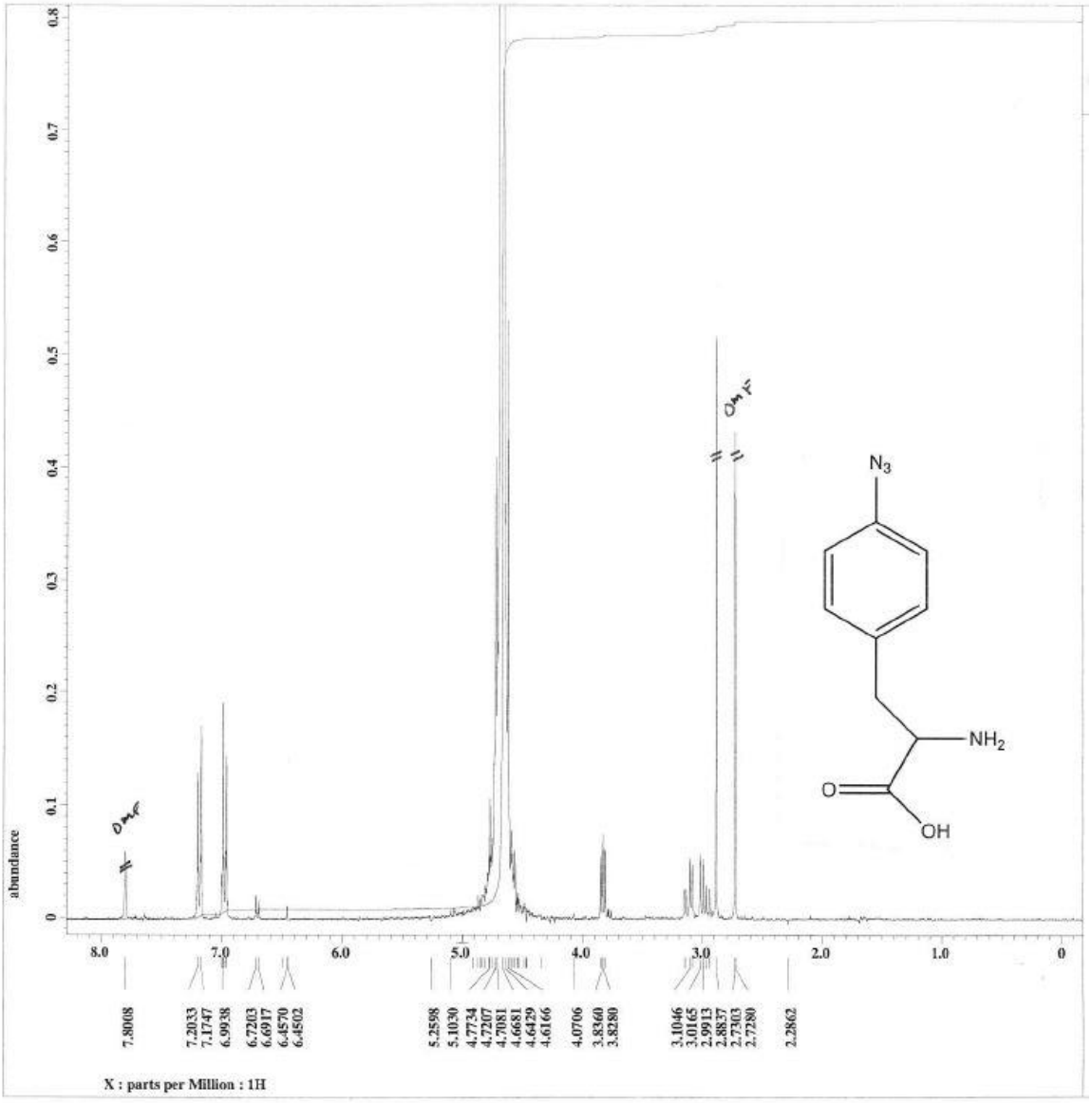
76. Davenport, E. P., Harris, D. F., Origanti, S. & Antony, E. Rad51 Nucleoprotein Filament Disassembly Captured Using Fluorescent Plasmodium falciparum SSB as a Reporter for Single-Stranded DNA. *PLOS ONE* **11,** e0159242 (2016).

77. Marini, V. & Krejci, L. Unwinding of synthetic replication and recombination substrates by Srs2. *DNA Repair* **11,** 789–798 (2012).

78. Meiners, M. J., Tahmaseb, K. & Matson, S. W. The UvrD303 Hyper-helicase Exhibits Increased Processivity. *J. Biol. Chem.* **289,** 17100–17110 (2014).

# APPENDICES

## Supplemental Figures




**Supplemental Figure 1 | 4AZP NMR.** Displayed above is the C13 NMR of our product, 4-azido-l-phenylalanine (4AZP). Some DMF is still seen with the product, otherwise the sample exhibits high purity.